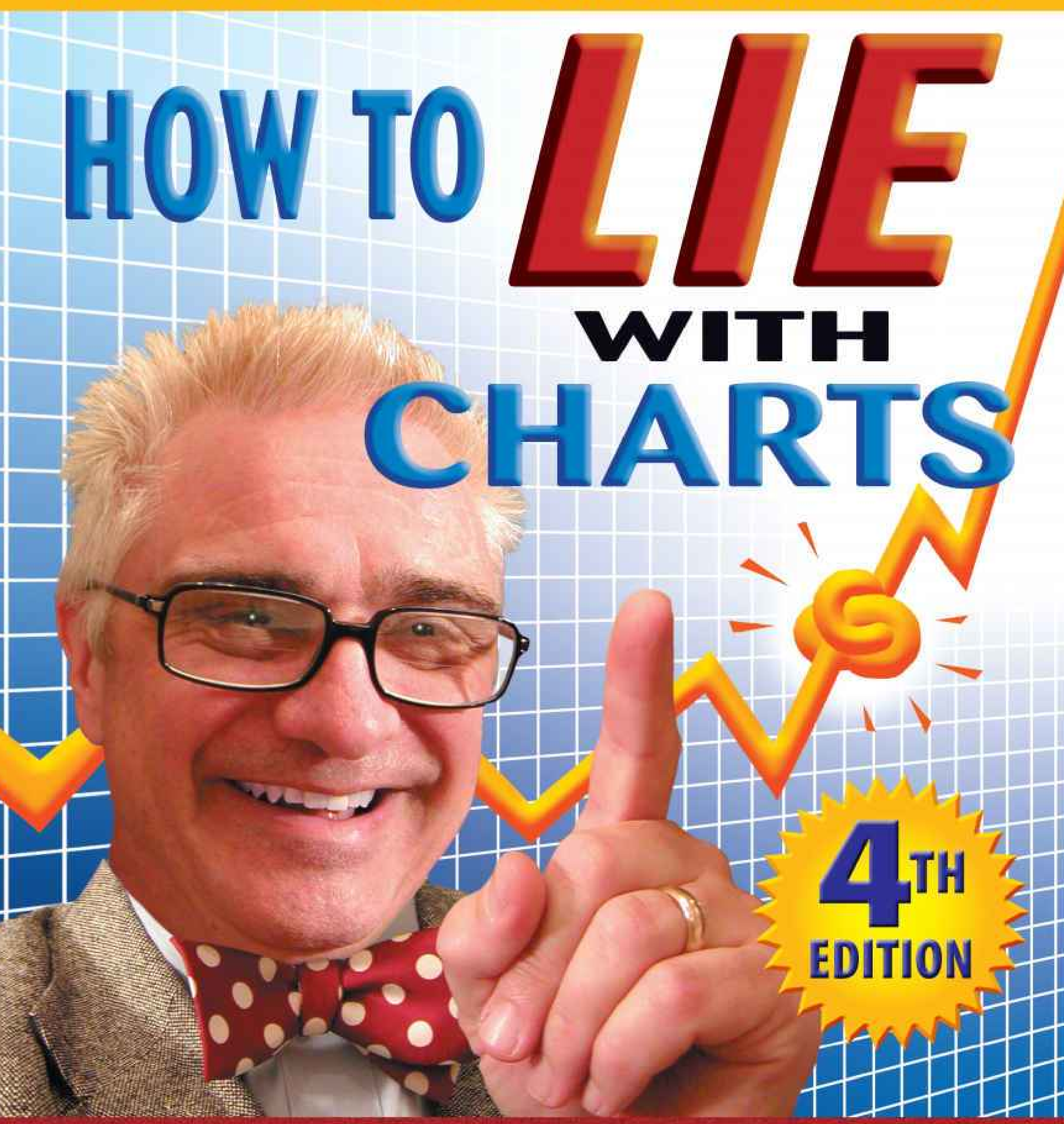
NEW CHAPTERS:
- FAKE NEWS IN SOCIAL MEDIA
- MISUSE OF METADATA AND BIG DATA



Gerald Everett Jones



How to Lie with Charts

Fourth Edition

Gerald Everett Jones



LaPuerta

Santa Monica, California

Developmental Editor: Richard Mills

Editors: Kristen Vanberg-Wolff, Thomas Page

Cover and Interior Designer: Gary Palmatier, Ideas to Images

Indexers: Ted Laux, Robin Bridge

ISBN: 978-0-9965438-6-6

Library of Congress Control Number: 2018911446

Contact the author or publisher: bookstore@lapuerta.tv

LaPuerta is an imprint of

La Puerta Productions

www.lapuerta.tv

This book is for ARTIS and my fellow wizards Roy, Ted, Joan, Charlene, Cindy, Sue, Maurice, Ed, Terry, and Tom.

Acknowledgments

MANY **dedicated and talented people** worked diligently to bring you this book, and I'd like to give some of them special mention here.

The credit for seeing virtue in its concept and the courage for advancing its cause goes to Richard Mills, developmental editor. I owe Richard a further note of thanks for taking me up on my recommendation for a book designer, my friend and colleague Gary Palmatier (Ideas to Images, Santa Rosa, CA). As far as I'm concerned, Gary and his associate Robaire Ream have outdone themselves this time, and I congratulate them for another highly professional job.

In the first edition (published by Sybex), a light but firm touch at the author control panel was exercised with wisdom by editor Kris Vanberg-Wolff. The indexer was Ted Laux. The third edition was edited by Thomas Page and re-indexed by Robin Bridge.

My colleague Matt Wagner has never sought recognition in my books, but he deserves my gratitude. He's been an able counselor and a sincere friend.

I've enjoyed the support and patient encouragement of Georja Umano, world-class actress, animal advocate, and life partner, whose biggest fan couldn't be in the audience some nights because he was shacked up with a laptop.

The notion of "Don't Show the Don't Knows" described in the Introduction and in Chapter 2 should be attributed to James Rothman and was kindly suggested to me by his son Edward.

The image of the motorcyclist in Figure 3.14 is © 1993 Management Graphics Ltd. (Reading, Berks, UK), used by permission. The chart is adapted from an actual graphic I saw in the sales brochure of a company I will decline to name. The photograph used in Figure 14.2 was provided by Gary Palmatier.

I am indebted to Yoshi Hatakeyama, CEO and Chairman of Zacta Engineering Corp. (Kawasaki City, Japan), for his advice and counsel on my references to Japanese culture in Chapter 5.

Thanks to Peter Nathan for suggesting the unconventional checkbook example in Chapter 9. Fred Gallegos, computer auditing expert, first introduced me to the worry of worksheets in his book *Audit and Control of Information Systems.* Fred's personal example convinces me that at least some G-men still wear white hats.

Andrew Hatoff provided information on color theory that was valuable in my preparing Chapter 14. Si Becker, past Director of Engineering for the Society of Motion Picture and Television Engineers, was the lab expert I consulted in the anecdote at the conclusion of the chapter on color. I owe much of my understanding of the subject of color to consulting engineer John Cool.

Much of the material in this book is based on software development work I did in the early 1980s as supervisor of the ARTIS project at Color Terminals International, a division of Creative Technologies, Inc.

I appreciate the hard work and diligent efforts of my foreign rights agents Neil Gudovitz and Kimberly Brabec of Waterside Productions. Thanks to Neil, you can read this book in Italian, Japanese, or in Complex Chinese.

I'm grateful to the reviewers who contributed helpful notes that have been incorporated in this edition. They include: Lee Higbee, Dale E. Schreiner, Kai-Mikael Jää-Aro, R. Kelly Wagner, Mike Broderick, Armchairinterviews.com, wiredweird, and Distance Cyclist. Walter Samuel sent me detailed comments, including quite a few suggestions and corrections that I adopted.

My heartfelt appreciation goes out to fans Robin Wolfson of Datastep Development, Betty C. Jung, MPH RN MCHES, and Oktay Özdemir—who from their raves seem to share a passion for truth-telling in business graphics.

I first presented the material on technical analysis in seminars I gave for the Association of Proposal Management Professionals. I developed the content on financial proofreading as part of my consulting engagement with the Credit & Collections Department of The Walt Disney Company. In addition to the overview seminar titled the same as this book, these supplemental seminars are available on DVD from LaPuerta as *Technical Analysis Basics* and *Financial Proofreading*.

The names of persons and organizations, as well as the data, used in examples and charts in this book are fictional, except where statistical data was derived from sources in the public domain. Deceptive techniques and mistakes shown in the charts are intended for instructional purposes only and do not describe or depict the work products of any actual person or organization.

Profuse thanks to my colleague and longtime fan of this book Prof. Betty C. Jung, who performed a much-needed reality check on the new chapters for the fourth edition.

Anyone who uses the tips described herein to lie deliberately with charts or social-media posts deserves an audience full of people who also read this book and take its precautions seriously!

Introduction: How to Lie with PowerPoint

A publisher once told me that I would sell 10,000 more copies of this book if I were to put the trademarked name of the popular Microsoft slide presentation software in the title—perhaps even more if we were to point a finger at Excel.

And a well-intentioned reviewer of the first edition wrote: “A better title would be *How to Avoid Being Fooled by Charts and How to Prepare Clear Informative Ones.*”

Hmmm. Honest, if not catchy.

Then a newspaper reviewer took me to task for teaching people how to lie. But he should have accused me of theft. I borrowed the title from Darrell Huff’s *How to Lie with Statistics,* which has endured mightily since it appeared in 1954. (And then Mark Monmonier did *How to Lie with Maps*.)

This LaPuerta publication is the fourth edition of *How to Lie with Charts,* which was first published by Sybex in 1995. Despite the fact that the first edition had a few embarrassing errors (yes, lied in its own way), it has become something of a classic among math teachers and business communicators. I assumed this reputation had preceded me when, years later, I met one of the new Sybex editors at an authors’ conference. As he shook my hand, he commented that he admired the book. Swelled with pride, I asked him if perhaps one of the senior editors at the publishing house had recommended it to him. “No,” he replied. “We studied it in school.” Can I be that old? Would you think I were lying if I claimed, like Mozart, to have turned pro at age four?

What's the Problem?

Are you worried about using PowerPoint? You should be. However, you can't beat the rap by blaming the computer.

Oh, don't get me wrong—the ability to generate charts and graphs at a click *does* compound the crime. Why? Because if you had to laboriously lay out your drawings using the ancient tools of graph paper, ruler, and pencil—you'd have time to *think* about what you were doing. But when the process is automated, many of the crucial design decisions happen behind the screen.

Isn't it marvelous digital magic? Sure. But when did you ever hear of a magician who was honest?

What, again, *is* the problem, exactly?

If you are using PowerPoint or Excel or any other computer application to make charts for an audience—it is almost certain that you are lying to them in some way.

So I don't need to teach you how to lie. You're doing it already.

But the problem is more serious than that. You don't know you're doing it. And neither does your audience. So you're getting away with it—until a manager or an investor or a sales prospect makes a bad decision based on the information that you were so helpful to provide.

But What About PowerPoint?

The main focus of this book is on the principles of visual communication. It's about careful thinking and clear expression. Computer technology is a side issue.

If you're still worried about using PowerPoint, I should back up my claims. By way of introduction to the underlying reasons, here are just

three ways that you and your computer helpmate can make downright unreliable charts.

1. Start by Loving Excel

And who doesn't? Life without spreadsheet software would be like an eternal morning without coffee. PowerPoint's Datasheet window is an embedded Excel object. You can also import data from external sheets and databases.

Your chart is no more accurate than the numbers from which it's generated. We're now living at a time when many young financial managers never did any manual calculation after high school (and maybe not even then). Unfortunately, students and managers alike are more concerned with getting results quickly than questioning the underlying logic of their sheets. And even if the logic is sound, inaccuracies from truncation or rounding in formulas or as a result of number formatting can be multiplied with every update (especially when a formula uses the results of another formula). It's all the more likely in a sheet that gets updated frequently—daily reports, say. Through iteration, margins of error can become significant—enough to matter.

The remedy? Learn to analyze the logic in complex sheets. And apply tests of transactions before you use the sheet on live data to generate charts. Oh, yeah—everybody knows that. But hardly anyone does it—life's too short.

2. Make Everything 3D

When you get around to selecting a chart type in PowerPoint, if you want to mislead or distort, choose dimensional pies or bars. It's virtually impossible to make a 3D chart that doesn't give a wrong impression.

In a 3D pie chart, the slice with the thickest edge will always seem larger than its true percentage. (Chapter 2 will tell you why.)

In a 3D bar chart, it's a challenge to find a rotation at which the short bars are still visible but the tall bars aren't distorted by the perspective.

Judging the heights of bars in a perspective view is a real challenge. So your audience can't reliably estimate values (do visual take-offs) from your dimensional bars. (See Chapter 4.)

Unfortunately, 3D bars are now the staple of the wildly popular "dashboard" displays of executive information systems. Managers are actually making operational and strategic decisions based on this type of reporting. It's scary.

3. Heed Rothman's Rule: "Don't Show the Don't-Knows"

I named this rule for British economist James Rothman, who preaches vehemently against following it. What he means is, in a data series, if you have some leftovers you're tempted to label "Miscellaneous," "All Others," or "Unknown"—it's so much tidier just to leave them off. Showing them is not only messy but it also invites questions about why you can't be more specific. In a bar chart, it's a negligent practice. But in any pie chart, it's an outright lie. A pie chart, by definition, *must* show 100 percent of the data set. (Read more in Chapters 3 and 4.)

Furthermore, if you are unsure of the value of the Unknowns, the relationship among all known items is almost useless. Since each pie slice is some percentage of the total, none of their sizes will be reliable if you don't know what the total is! Yes, you know how the knowns relate to one another, but you have no clue how significant any of them is in relation to the whole.

Be warned: There is no built-in control in PowerPoint or in Excel or in any other computer charting application that will force you—much less, remind you—to account for 100 percent of the real-world data in a pie, or in any other, chart.

If you're ready to swear you'd never resort to this low trickery, think about what programmers call the "unk-unks," the *unknown* unknowns. Whatever the situation is that you're trying to depict, I'll bet you a trip to the snack bar there's something about it you don't know. If you think there isn't, you're lying to yourself.

What Makes Me So Sure?

My background in business graphics has spanned the era from mainframes to PDAs, from VHF-TV to cellphone mobisodes. Much of the material in this book is based on software development work I did in the early 1980s as supervisor of the ARTIS project at Color Terminals International, a division of Creative Technologies, Inc. It was the era of the first IBM PC and the Apple II. ARTIS was based on a full-color Z80 Cromemco microcomputer, designed as a user-friendly desktop design station for the room-sized Genigraphics slide making system (which was built around the DEC PDP 11 minicomputer). The output was 35mm color slides, each of which could take several minutes to render and record on a 2,000-line CRT film recorder, followed by a time-consuming bath in E-4 photo-developer chemicals and a dizzying run though a machine that cut and mounted each film chip in a small plastic frame.

ARTIS was an impressive technical achievement but a commercial failure. Its potential competitors did not fare much better. However, eventually the MS-DOS application Harvard Graphics caught on, with a look and feel much like the best of its forebears, including ARTIS. Lotus Development then released Freelance Graphics. After these products migrated to Windows, Microsoft won virtually the entire market for itself by deciding to bundle PowerPoint with Word and Excel in its incredibly successful Office suite.

Today, PowerPoint is the standard. However, in my opinion, it's not the standard because it's the best. Freelance Graphics for Windows was probably the most robust and flexible business graphics application ever developed. PowerPoint is the standard by default, not because it was better software but because Microsoft dominated the marketplace and bundled it in Office.

While PowerPoint was conquering the desktop, a revolution in display technology virtually eliminated color slides. Back in the early 1980s, Eastman Kodak was busy estimating how many millions upon millions of slides would be consumed by business presenters by the turn of the century. The big debate was about whether computer-generated

slides were sharp enough to meet expectations of quality in corporate boardrooms, auditoriums, and exhibit halls. What no one figured on was the incredible convenience and portability of the laptop computer and the ease of plug-and-go video display. Graphics displayed on the computer screens of the ’80s and early ’90s had a much more jagged look than did computer-generated slides, but convenience—and the zeal for instantaneous updates—won out over visual quality. Nowadays, digital video projection rivals the look of 35mm theatrical film, but the veterans still complain about quality. (Film is dead. Get over it.)

Actually, I’ve been harping about this stuff for a long while. Portions of this book first appeared in a series of magazine articles I did for *Computer Graphics World* magazine (1983-1986) and in my books *Looking Good with Harvard Graphics* (ScottForesman, 1991), *Murphy’s Laws of Excel* (Sybex, 1993), *A Guided Tour of Excel* (Sybex, 1994), *Font Secrets and Solutions* (Sybex, 1994), *Harvard Graphics: The Art of Presentation* (Prima, 1992) and *Freelance Graphics: The Art of Presentation* (Prima, 1993). In 2004, Cliff Atkinson (author of *Beyond Bullet Points*) invited me to submit the article “How to Lie with PowerPoint Charts” to the Accent Technologies website, and the gist of that material has made its way into this introduction.

What Else Is New?

When the first Sybex edition went out of print in 2000, it was republished as a print-on-demand book by iUniverse. That kept the book on the market and selling steadily, even though it hadn’t been updated. A drawback was that, in those early days of POD, master files for digital printing were created by optically scanning the pages of the existing book. As a result the print quality was rather fuzzy, and the reproduction of the illustrations, particularly the gray halftones, was poor.

For the second edition, we made many corrections, updates, and clarifications. Readers will find substantial changes in Chapter 14, where the original emphasis on color design for different output media (such as slide film and overhead transparencies) has now shifted to

consideration of room lighting conditions for video display or projection (dark versus bright rooms). The discussion of color palettes, which follows closely the way color was implemented in Freelance Graphics, might seem dated since PowerPoint doesn't give the user as much access to these choices. However, on closer examination, the principle of palette-driven color schemes is essentially the same as in the Cascading Style Sheets (CSS) that are so much a part of Web design, especially when used in conjunction with "format-agnostic" XML.

We aimed the third edition squarely at investors and financial analysts. New material on technical analysis, dashboard displays, and financial proofreading was added as Chapters 10, 11, and 13 respectively. To a large extent, I developed this material in response to questions I got when I gave seminars to business groups based on the original book.

The new material in the fourth edition addresses the lamentable reality that fake news is no joke—and yesterday's weather report is not a reliable basis for metadata analysis.

As we venture into the gray areas in the following chapters, remember that deliberate lying isn't my concern, nor do I mean for it to be yours. As a trained investigator once told me, financial loss from unintentional error is easily 10 times greater—perhaps even 100 times greater—than the cost of fraud. As reviewer Lee Higbie put it, "Don't assume conspiracy when stupidity will explain it."

So be guided by the Golden Rule of the Information Age:

DON'T SHOW UNTO OTHERS
WHAT YOU DON'T WANT
SHOWN UNTO YOU.

1

The Numbers Don't Lie

Or Do They?

NO doubt you've heard a colleague defend a business decision with the sweeping declaration: *The numbers don't lie!* Surely, a conclusion based on solid data *must* be reliable. Now, if you swallow this little bit of wisdom whole, you might also be tempted to believe that charts based on those same numbers must also show the truth, the whole truth, and nothing but the truth.

Oh, if only our world were so simple!

This book can teach you how to lie with charts. Or, if you don't tell lies, you can learn how to defend yourself against the shameless people who do. In this chapter, I'll explain how you or anyone can begin by lying with the numbers—that none-too-solid foundation on which your charts are built.

Have I given you the wrong impression? Please don't think that I want to make it easier for you to be dishonest! Your sense of business ethics is not at issue here.



Here's the major premise of this little book: Learning how to lie with charts will help you spot deceptive tricks, whether in someone else's charts or in your own. Where you go from there is up to you!

A chart is a graphic presentation of a set of numbers. A mathematician would refer to them as a *data set,* or *data series.* The set of numbers represents a collection of measurements. These measurements

typically describe how some *thing*—say, the growth of a company or a wart on your toe—is changing over time. Another type of chart can show the same measurement taken for several things at some "snapshot" in time.

All charts distort the truth, if by *truth* you mean *reality*— the physical world. Charting is a type of *data reduction* whereby the complex world impinging on our senses can be made simpler, and thus easier to understand. Data reduction isn't necessarily a bad thing—perhaps you've heard that *less is more*—but you have to be careful doing it.

The opportunities to distort reality don't begin when you reduce a set of data to a chart, however. There can be all kinds of distortions in the numbers themselves. In fact, the apparently simple act of *counting* can produce serious errors.

Counting Is Simple—Isn't It?

$$$$INVESTOR'S TIP$$$$

A real-time chart of stock market results would make your head swim. But a five-year performance chart may smooth over some recent rough spots you need to see. It's all about what the analysts call the *level of abstraction.*

We perceive reality by measuring its effects on our senses. For example, we can count the things we see. However, just the act of counting—without doing any higher math—requires you to interpret, and possibly to distort, reality.

Okay, you walk into the produce section of your local market and start counting pieces of fruit. When your task is done, you hand me a number:



This number is a piece of data. Like garden produce, the data is *raw.* It is *unprocessed*. Raw data is usually, well, downright indigestible. Digesting (understanding) it involves that fearful term —*data processing*!

Now, you can't avoid the processed stuff. You need it just to survive. Face it, reality is messy. It can involve lots and lots of numbers—computers full of them. And handling all those numbers can be inconvenient. They arrive all jumbled, just the way reality generates them—not in nice, neat rows and columns, much less in pretty charts.

By the way, what did you mean by that number you gave me?

19 WHAT?

$$$$INVESTOR'S TIP$$$$

It can be all apples and oranges in reports that intermix audited and unaudited results, different fiscal year-ends, and partial or estimated current years.

Are we talking apples or oranges here? Even in the act of collecting the data, you must impose your own notions on reality: Okay, you counted apples. Did you count just the red apples or did you include the yellow ones, too?



Right away, I can see a problem with *categories*. Just trying to count things means we have to make decisions about categories—which things to include and which to exclude from the count. For example, how about that half-rotten apple over there? Surely, you didn't count that. You can't really call it a piece of fruit. After all, it's not a whole piece. And, if nobody wants to eat it, would you call it fruit? Or just garbage?

My Apology to the Citizens of Rome

The word *data* is the plural of the Latin *datum,* meaning *fact.* So it would be perfectly correct to say, "The data are in Harvard Yard." However, even careful lexicographers admit that data can be a collective noun (referring to a group containing more than one datum), in which case it would still be impeccable usage to say, "I went to Georgia Tech, and the data is wherever I say it is." In this book, I follow the crowd in saying "data is," but I will defend not quite to the death your right to say it either way.

Is There Any Truth in Labeling?

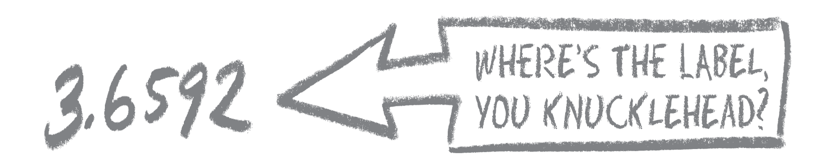
The categories of things can be indicated by labels, which are as necessary to reliable data as the numbers, or *values.* As a rule, wherever you see a number, you should find a label nearby.





In a chart or in a printed report, if you see a number without a label, stop right there. Demand a label, and don't take "None" for an answer. You have a right to know what you're dealing with, after all. There's more advice about labels in Chapter 12, "Words for Your Charts, Or Is There Any Truth In Labeling?"

My high school chemistry teacher Mr. McNeill was a fanatic about labeling. On exams, if you left a label off an answer, he gave you no credit for it at all, even if you got the number right.



McNeill preached that a number without a label is meaningless. (Well, that's not strictly true. It has a meaning in math, but that was Miss Bowers's class.) By itself, a number is not useful.

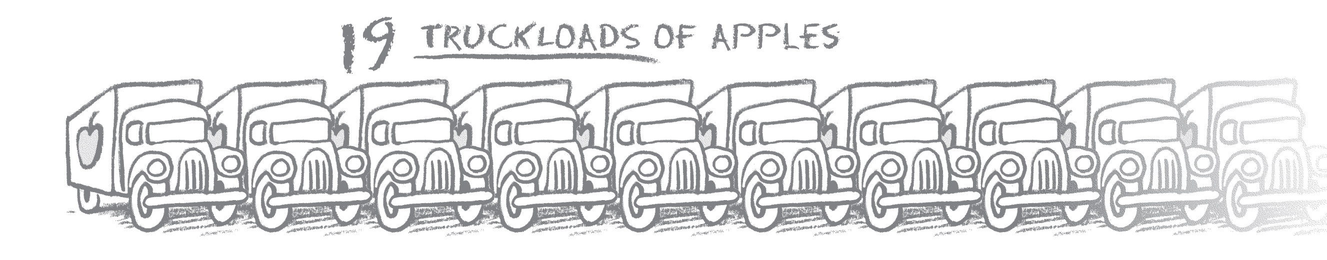
Often a label must have two parts—a unit of measure and a description of the thing being measured:

3.6592 LITERS OF H_2SO_4*

*SULFURIC ACID – CAREFUL!

Getting back to your apple-counting, did you count 19 apples or 19 pieces of fruit? Was it 19 pieces, 19 dozen, or 19 cases?

Add a label to a number, and you've really got something:





Leaving a label off an important number is a favorite trick of obfuscators and graphic deceivers. Worse, using a label that is only slightly incorrect is the mark of a truly skilled liar. For example, if I were trying to exaggerate a count, I might say 19 *crates* instead of 19 *boxes*. *Crate* seems bigger, more substantial, than *box*, giving the impression that it can hold more, even though there might be no difference in the physical capacities.

Another set of labels can relate this stuff to the things around it:

TOTAL INVENTORY:
19 CRATES OF APPLES

A naked number is one kind of raw data. It's not much use by itself. Add a label, and the data describes something in the real world. Relate that description to the things around it, and you have useful information. And, after all, getting useful information should be the goal of every honest chart maker!

What Is Information?

One definition of information is that it answers a question:



Funny about information. Because our brains find it useful, it often stimulates more questions than it answers:

- How many apples did we have yesterday?
- How many apples did we sell today?
- How many apples should we order for tomorrow?
- Who wants this half-rotten apple?

With the exception of the bothersome rotten-apple issue, one chart might be able to answer all these questions. A single chart might even describe an entire apple inventory at a glance, as shown in Figure 1.1.

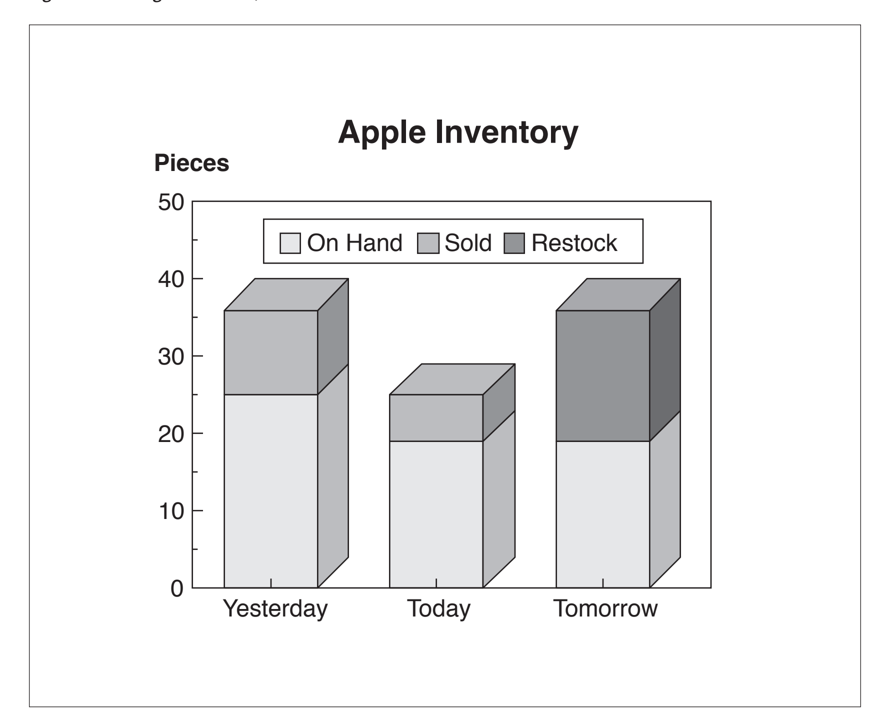
Data or Information?

Here's a familiar example:

DATA 8005551212
INFORMATION (800) 555-1212

The raw data is a string of digits, which might represent anything. To be useful, the data must be interpreted. The useful form—information—is obviously the second one. Each part of the phone number—the area code, the exchange, and the extension—acts as a label, or data category. The categories are the keys to interpreting the data.

Figure 1.1 *This snapshot of the inventory shows how many apples will have to be ordered today to replenish the stock to 36—yesterday's starting level. (Even this simple chart distorts reality, because you can't order apples by the piece. You have to buy them in case lots! Also, see Chapter 4 for cautions against using 3D bars.)*

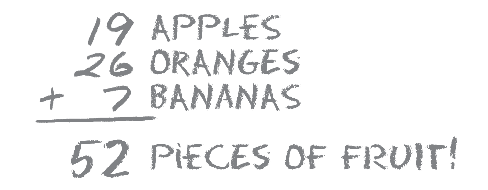


Even though the chart addresses all these questions—a real cornucopia of information—its data content is *reduced,* as compared to the raw data that was generated from counting apples. No matter how much you study the chart, you cannot know everything about the real produce bin in the store! For example, you don't know from the chart whether the yellow apples were counted along with the red ones, and you certainly don't know about the characteristics of individual apples. How many of them, for example, are bruised, blemished, or rotten?

To make decisions based on such a chart, you have to make some assumptions about the underlying data. Could I assume that you only counted salable apples? In the end, anyone who didn't do the counting would have to trust the person who made the chart.

Avoid Passing Summary Judgment!

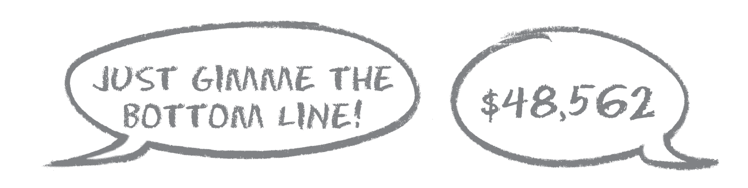
An everyday kind of data reduction is an arithmetic *summary:*



Who said you can't mix apples and oranges? Summarizing a set of numbers to produce a total—the infamous bottom line—reduces the complexity of reality. In this case, the reduction is down to a single number. The total is more understandable than a collection of raw data. It's tempting to work with totals because we can quickly get



Remember, though, that the process of data reduction obscures the underlying details, which may include important facts. For example, consider the business manager who wants the bottom line on monthly sales. She wants the total, and quickly:



Pleased with this number, the manager might move on to other, more pressing matters. However, as in horror films, danger often lurks just beneath the surface. Here is the raw data summarized by that total—results according to each salesperson on her staff:

THIS MONTH'S SALES
BY SALESPERSON

D. AVALOS	$45,281
P. NOONAN	0
A. ONO	2,919
F. BERG	362
TOTAL	$48,562

A deeper inspection of the data reveals that two of the salespeople haven't done much at all, and another had only modest results, while one of them has racked up most of the sales. What's the story here? Are the other three loafing, or have they been developing good prospects that might generate sales next month? Is that stellar performer overworked, perhaps pushing toward a nervous breakdown? The manager would certainly ask these questions if she looked closely at the data, but she could ignore a possibly precarious situation if her focus stays only on the bottom line.

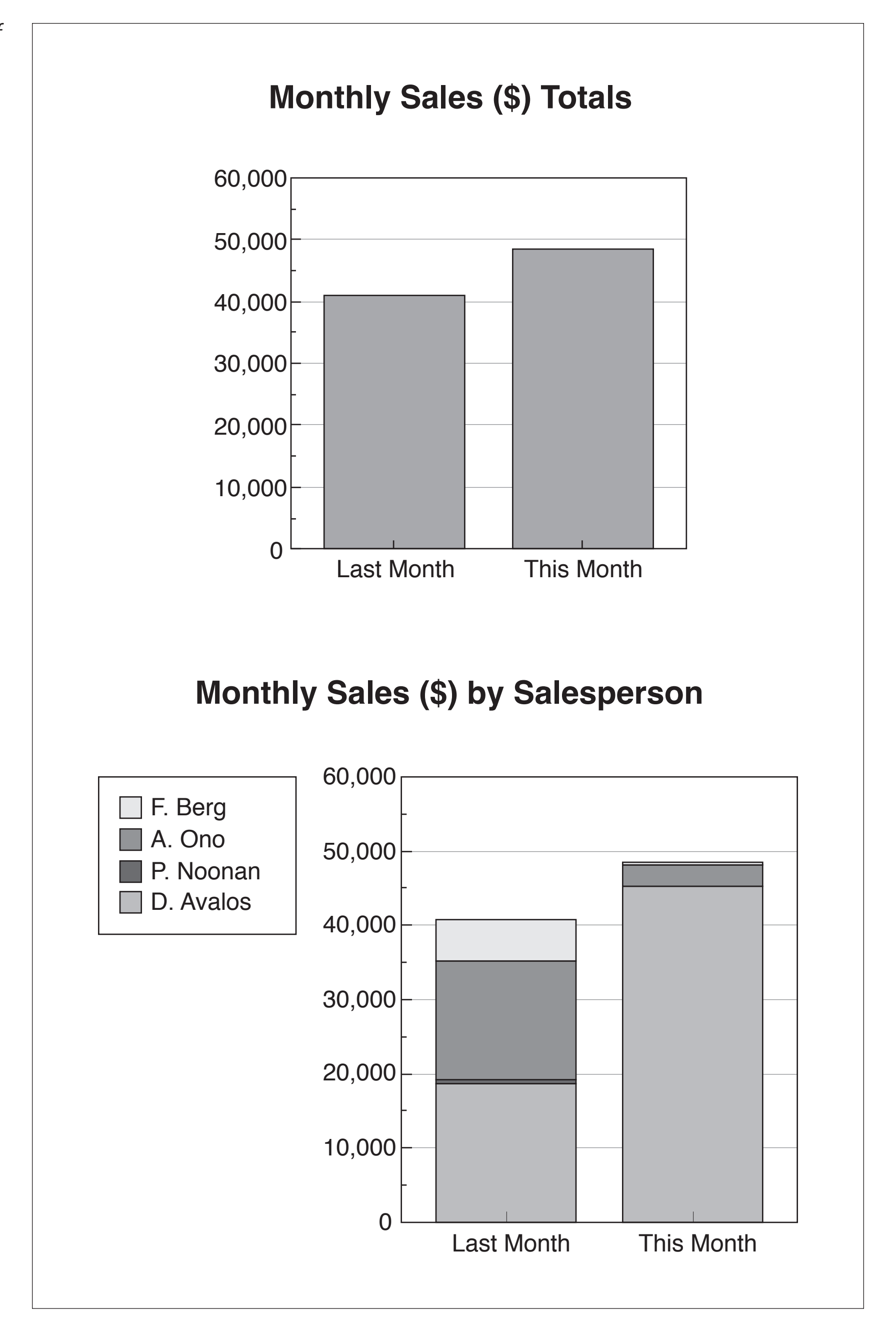
$$$$INVESTOR'S TIP$$$$

It's necessary and desirable to consolidate international results by converting to a common currency. But do remember that conversion rates change on a daily basis, so your chart should state it is only a snapshot on a specific date.

Compare the amount of information you get in the two charts in Figure 1.2. The top chart shows a bar graph of the bottom-line results—total sales. The bottom chart shows the same totals, with the addition of *segmented bars* to show individual contributions. The sales manager would do well to insist on seeing the data presented this way.

The tricks of data reduction go way beyond generating totals. Perhaps you've heard your colleague refer to "massaging" the data? Well, one way to get a good numeric rubdown is to start playing with your chart data as *averages.*

Figure 1.2 *Two views of sales results: The use of segmented bars reveals a hidden story.*



You Can't Trust Your Average Golfer!

A player's proficiency at games such as golf, bowling, and baseball is usually described as an average, another type of data reduction. If a golfer at your office says he shoots in the high 70s, he is *supposed* to be saying that, on a typical day, it takes him less than 80 strokes to knock a little white ball into 18 successive holes on a regulation fairway. The rookie liar, of course, adopts his low score from one exceptional game as his lifetime average. But if you want to lie to the pros, you're going to have to do better than that!

The average, or arithmetic *mean,* is the sum of the scores of previous games, divided by the number of games:

$$\frac{79 + 81 + 78 + 80 + 76 \text{ STROKES}}{5 \text{ GAMES}} = 78.8 \text{ AVERAGE STROKES PER GAME}$$

The average is meant to describe a typical, or usual, score over a player's history. Like summaries, averages sacrifice the underlying detail in favor of a single, more understandable number. And as in the bottom-line sales example, relying exclusively on an average can lead you to incorrect conclusions.

A big problem with using an average as an indicator of performance is that averages are reliable only if the underlying data values are consistent. For example, a golf average of 83 is a fair picture of a golfer who had the following history:

$$\frac{85 + 83 + 82 + 84 + 83 \text{ STROKES}}{5 \text{ GAMES}} = 83.4 \text{ AVERAGE STROKES PER GAME}$$

However, the same score—even though correct as an average—is a poor description of this player's performance:

$$\frac{74 + 73 + 72 + 100 + 98 \text{ STROKES}}{5 \text{ GAMES}} = 83.4 \text{ AVERAGE STROKES PER GAME}$$

This unfortunate player was shooting in the low 70s (a real pro), then slammed a car door on his hand on the way to the fourth game. As a result, his average is *skewed high*, distorted upward by the high scores he shot when he was recovering from his injury.

Similarly, a single average score won't tell you if a player is erratic—perhaps performing brilliantly on rare occasions, dismally the rest of the time.

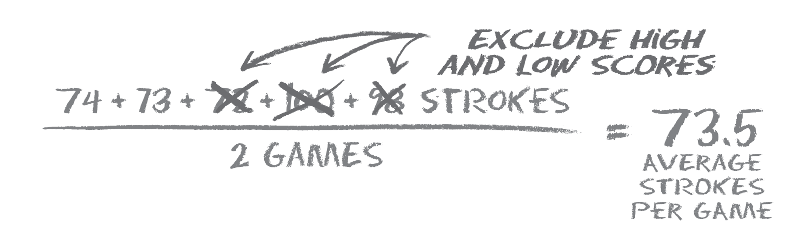


An average will be more reliable if: 1) there is a large number of individual data values, thus minimizing the effect of any one of them on the result, and 2) the data values are fairly consistent, or without wide variations, thus reducing the effects of unusual values on the result.

There are all kinds of ways of manipulating averages so that they are closer to your own idea of what's typical.

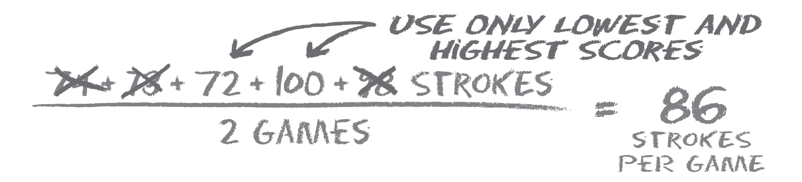
Throw Out the Highs and Lows

Assuming that very high or very low scores are too unusual, just disregard them and average the rest. A pitfall here is that the overall average will be based on fewer games, reducing the reliability of the average. To be fair about comparisons with other players, you could make a rule that you will throw out the same number of games for everybody, even for consistent players.



Use the Midpoint

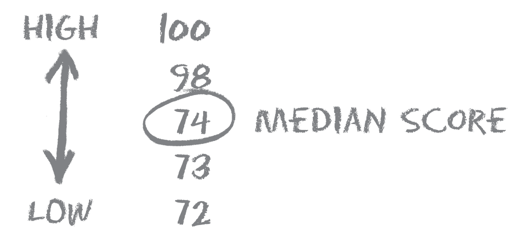
Take the average (arithmetic mean) of only two values, representing the high and the low, by adding them and dividing by 2. The result is the *midpoint*—not the overall average. The midpoint might be a better indicator if the data values are erratic.

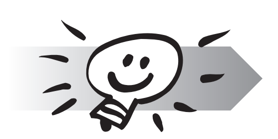


Use the Median

$$$$INVESTOR'S TIP$$$$
Why do you think political pollsters favor median scores on opinion surveys?

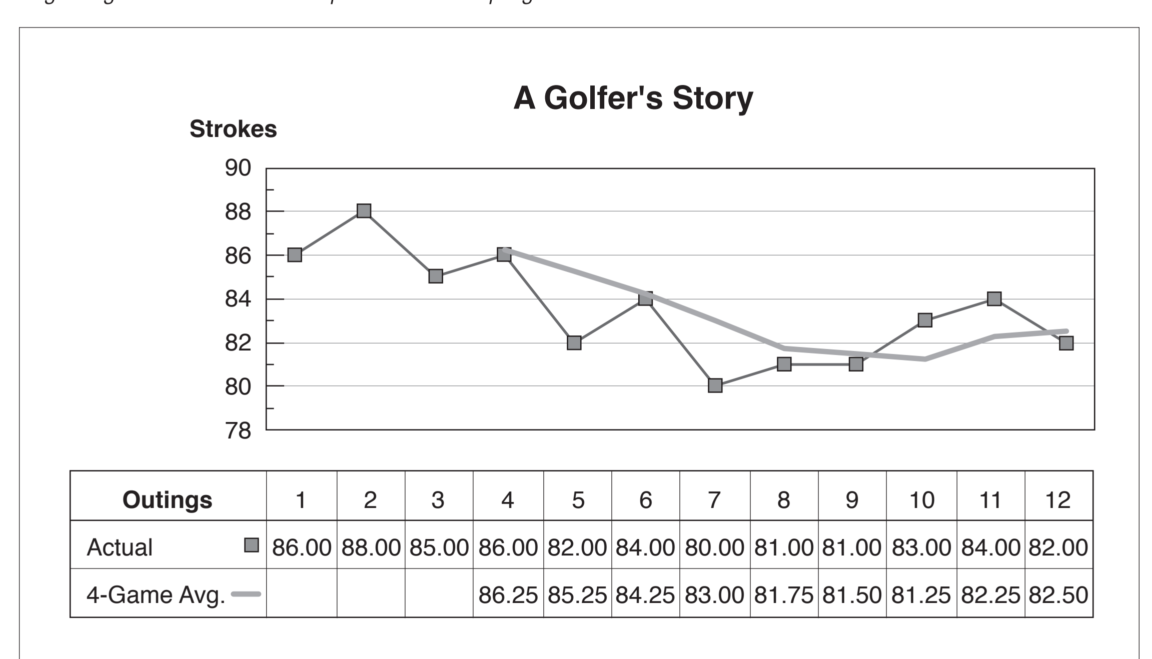
In a group of data values, the *median* is not an average, but one of the values that seems "normal." The median value in a set is located exactly in the middle *of the count*, where there are the same number of values above it as are below it. So, if you have a set of five scores, there will be two scores above the median and two scores below it.





If you have an even number of data values, you would take the midpoint of the *middle two* as the median. For example, in a set of four values, you would take the midpoint of the second and the third values, so that there are two values above and two below it. A related concept is the *mode,* which is simply the value that occurs most often in a set of values. The mode is therefore another "typical" value.

Figure 1.3 *The smoother line is the player's four-game moving average, which some people might regard as a better description of overall progress.*



Outings	1	2	3	4	5	6	7	8	9	10	11	12
Actual	86.00	88.00	85.00	86.00	82.00	84.00	80.00	81.00	81.00	83.00	84.00	82.00
4-Game Avg.				86.25	85.25	84.25	83.00	81.75	81.50	81.25	82.25	82.50

$$$$INVESTOR'S TIP$$$$

Stock market charts in the daily newspaper typically plot the 65-day moving average. The result is to show overall, longer-term trends in the market value of the stock in contrast to daily and weekly fluctuations.

Use a Moving Average

Averaging all your golf scores over your lifetime would be unfair. You don't want your recent scores to be distorted by your poor beginning scores, which should become less and less typical as you improve your game. Instead, a *moving average* might include only your most recent set of games—say, the last four. As each new game is included in your average, the oldest game drops off so that there are always just four games included. A chart of your lifetime performance, then, would show not only your actual scores, but also a series of four-game averages (see Figure 1.3).

Don't Count Blanks

Whenever you are working with averages, you can get into real trouble if there are any blanks, or gaps, in the series of data values. In the case

Will Those Lying Golfers Stop at Nothing?

If playing with their averages doesn't yield the right numbers, golfers have yet another way of manipulating their scores—the *handicap.* For a less-experienced player, a handicap is some number of strokes that always gets subtracted from that player's actual score. The golfers rationalize it this way: The score adjusted by the handicap compensates for widely different levels of skill so that it is possible to make comparisons with scores of more experienced golfers. So, here's an important lesson in business math: When none of the conventional methods produces the answer you want, *invent a new method!* (Be sure to give it a clever name that invites sympathy.)

---

of golf scores for several players, a blank might represent a day when one of the players didn't show up. However, a position that holds a zero is not a blank. In general, it is good practice to use zeros rather than blanks in a data set. When counting how many values will be averaged (the number you use as a divisor), *exclude the blanks* but *include any zeros*. Remember that zero values will distort the average unless the other values are near zero (0.1, say).

These are just a few of the many ways you can fudge—er, manipulate—data to reduce a set of values to one value that's typical, normal, simple, digestible, understandable—but nonetheless misleading. In the last few examples, I used various methods to get typical golf scores of 83.4, 73.5, 86, and 74 *for the same player history*! Each of these techniques is a perfectly valid type of data reduction commonly used in developing data for business charts. The particular technique employed is mostly a matter of judgment.

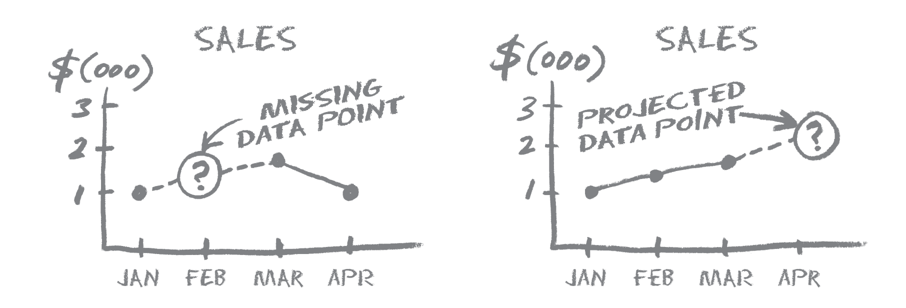
Which score do *you* think should be regarded as typical for the player with the injured hand?



Beware of obfuscators who use the term *average* loosely to refer to any of these very different techniques! (And besides the arithmetic mean, there's the weighted mean and the geometric mean.)

Trendy Thinking Can Be Risky

Just as dangerous as data reduction is any attempt to *expand* a set of data, either by guessing in-between values *(interpolation)* or by identifying trends and projecting future values *(extrapolation)*.



Using charts to identify trends is probably the slickest of the expert liar's tricks, and this stuff can get ambitious, mathwise. I spill my guts on this subject in Chapter 8, "Trends—Or, A Conservative's Guide to Trendy Thinking."

The Bottom Line on Data Reduction

$$$$INVESTOR'S TIP$$$$

Although it's a bit of an exaggeration, you may be better off regarding each chart as expressing the opinions of the chartmaker rather than objective findings.

I admit that this chapter offers many problems, with few solutions. But I want you to develop a healthy sense of caution about drawing conclusions from charts and chart data before we get down to the graphic details in the chapters that follow. My intention is to begin by making you sniff the air, so to speak, for those rotten apples.

As examples in the rest of this book should demonstrate, charts aren't simply neutral vehicles of communication. They don't just show data. Charts are ways of reducing *and interpreting* data to accompany a business message.



You will learn that there is at least one interpretation of the data in any chart, *even if its designer did not put it there deliberately*.

Debaters and trial lawyers make their points through persuasive argument. Think of a chart as an argument made visual.

The trick, if you'd call it that, is to make a chart's interpretation of the data *support and reinforce* your business message. That's not necessarily lying. It's just good communication.

2

Pies

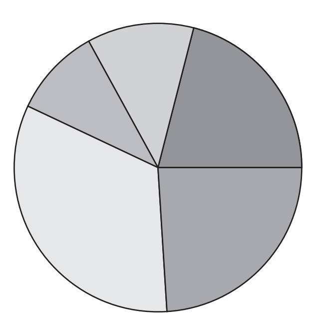
Or, When in Doubt, Throw One at Them!

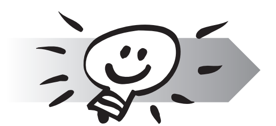
EVERYONE **wants a piece of the pie.** The *pie chart* is easy to make, easy to understand. Right?

The familiar pie chart is the most overused, misused, and sometimes downright useless trick in the presenter's repertoire. Liars love them, and in this chapter I'll show you why. You'll also find out how to throw one at your audience—if you really must—without making an awful mess.

The "Why" of Pies

This type of chart is, yes, easy as pie—a circle cut into slices:

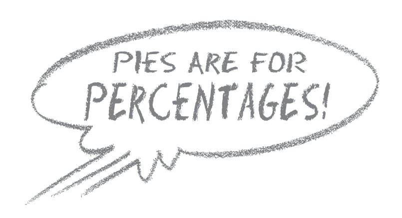




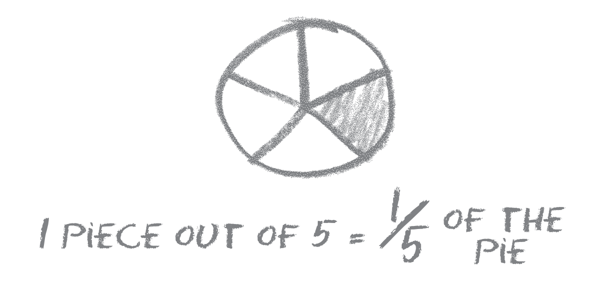
Other generic terms for the slices of a pie are *wedges, segments, cuts,* or *pieces.* The correct term in geometry for the slices of a circle, however, is *sectors*. (Okay, who cares?)

> **$$$$INVESTOR'S TIP$$$$**
> Be immediately suspicious of any pie chart that shows monetary values. Again, pies are for percentages. Showing the dollars as well might be honest, but the important message lies only in the ratios.

There's a simple rule for using pie charts correctly:



Instead of percentages, you can use a pie chart to show any proportional relationship between a slice and the whole pie. Such a relationship can also be expressed as a fraction, a ratio, or a decimal. For example, "This piece is one-fifth of the pie"



can be written

AS A FRACTION: 1/5
AS A RATIO: 1:5
AS A DECIMAL: 0.2

Multiply the fraction by 100—the number of percent in a whole pie—and you get its percentage of the whole:

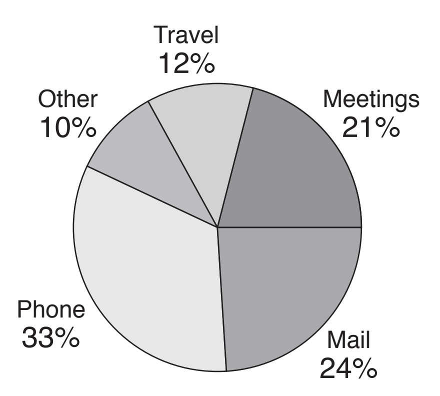
$$\frac{1 \text{ PIECE}}{5 \text{ PIECES}} \times 100\% \text{ (WHOLE PIE)} = 20\% \text{ OF PIE}$$

Showing any of these relationships—percentages, fractions, ratios, or decimals—is the purpose of pies. But you will be inviting liars to your table if you serve up your pies as anything else!

The trouble starts when you begin to think of pies in terms of *amounts*. A symptom that you've slipped into this mode of thinking is showing amounts in the slice labels. But, you ask, what's wrong with providing such potentially helpful information?

Percent *of What?*

In the first chapter, I strongly advise you to put labels on numbers. A number without a label, so I say, is useless. Now, pie slices should have labels, too. You and the audience have a right to know what you're dealing with.



Okay, fine. It's clear what each of the slices here represents—some percentage for each labeled slice. But, you all want to know,

- What does the whole pie represent?
- What is *its* label?
- These slices are percentages *of what?*

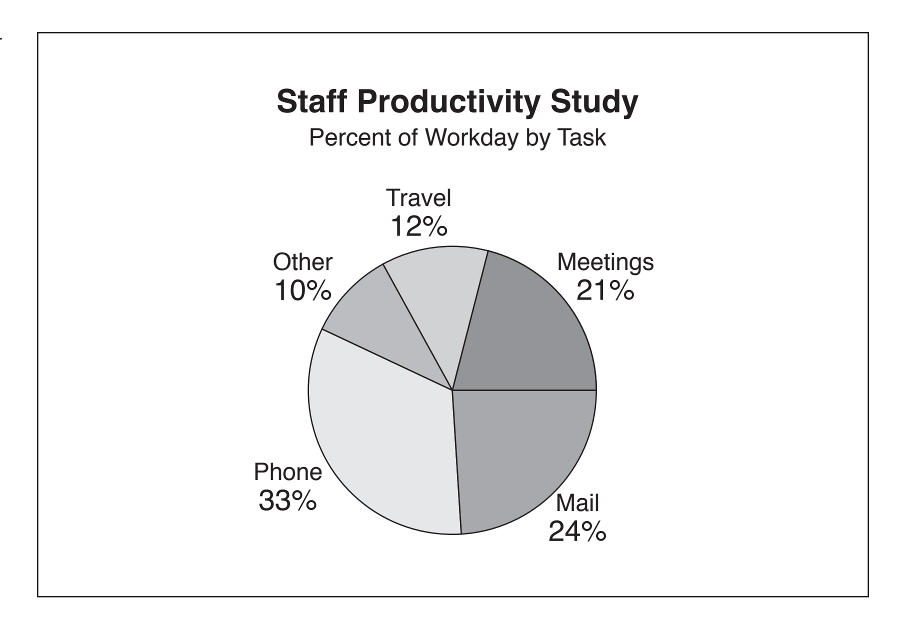
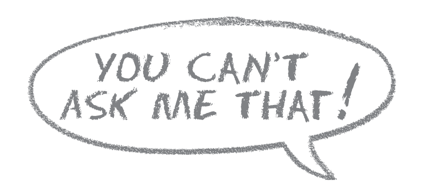


Figure 2.1 *The label for the whole pie should be the title of the chart.*

The label for the whole pie should be the title of the chart, as shown in Figure 2.1. In this example, each slice represents some portion of the total hours in a workday. The chart, then, is meant to show how the members of a sales staff are dividing their time among required tasks.

The whole pie represents the amount of time in a person's workday. So, maybe you now want to know, "How many hours are there in a workday?" Or, "How much is the total pie?"



In a pie chart, the *amount* of the whole is hidden from you! This secret feature of the pie chart can be a dirty little trick —or a redeeming virtue, depending on how you use it.

The purpose of a pie chart is to focus the attention of the audience on percentages—on the *relative* sizes of the slices. The actual values of

the slices and of the whole pie should not be important. The answer to "percentage *of what*?" should always be "of the whole pie." If you need to know more about the *value* of the whole, *you shouldn't be using a pie at all!*

Look back at the time study in Figure 2.1. Why can't you ask, "How many hours in a workday?" You might assume that a workday is eight hours long, but, in this case, you shouldn't. This chart is just as valid for the part-time staff who work just two hours as it is for the overachievers who put in twelve!

The intent of this chart is to support a discussion about how much time salespeople spend traveling to meetings in relation to all their available time—the whole pie—and in relation to the meetings themselves. The total amount of time—the number of hours each day—is not important to the intended message.

What is wrong, then, with the chart in Figure 2.2?

The pie charts in Figures 2.1 and 2.2 are identical, except for the slice labels. The first chart shows each slice, or type of task, as a percentage of a total workday. The second shows each slice as an *average number of hours* in a workday.

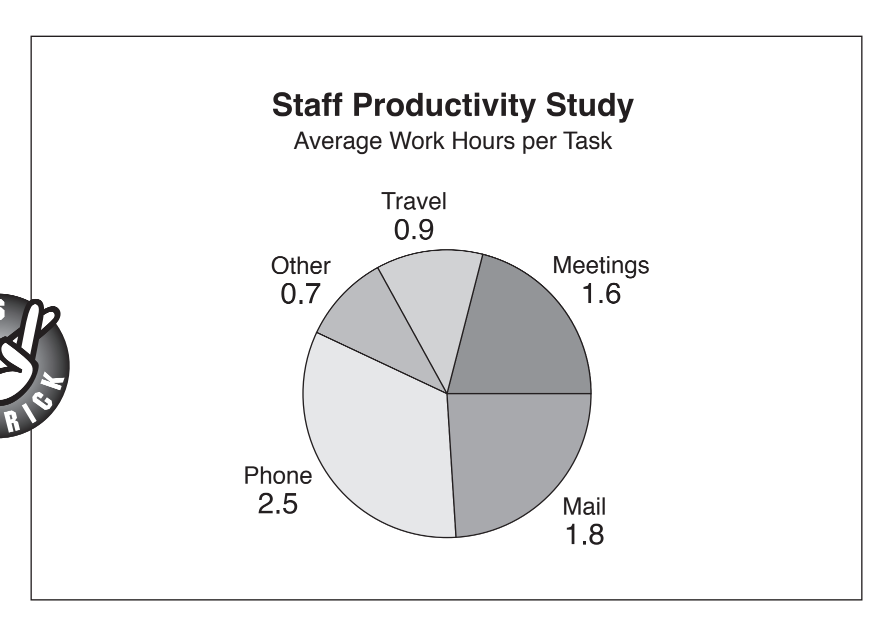


Figure 2.2 *This pie chart shows values instead of percentages for the pie slices. It gives the possibly incorrect impression that most people on the staff work 7.5-hour days.*

When actual values are given for pie slices, the audience will be tempted to add them to find out how much the total is. (If they don't do this, they will at least have some mental impression of the total, even if they are not fully aware of it.) In the case of the time study, the total number of hours is 7.5—not 8. Again, these are average hours. The part-time people have been included with the overachievers and everybody else. In fact, the majority of the staff works full 8-hour days. The length of the workday in the chart has been skewed by including the others in the average.



Look back at Chapter 1 to see how deceptive averages can be!

The presenter's intention is to recommend to top management that salespeople try spending more time on the phone and in electronic conferencing and less time traveling and in face-to-face meetings. However, the presenter could get easily sidetracked by irrelevant questions if the top managers become more worried about how the staff is spending the other half-hour each day—the 0.5 hour that seems unaccounted for in the daily average of 7.5!

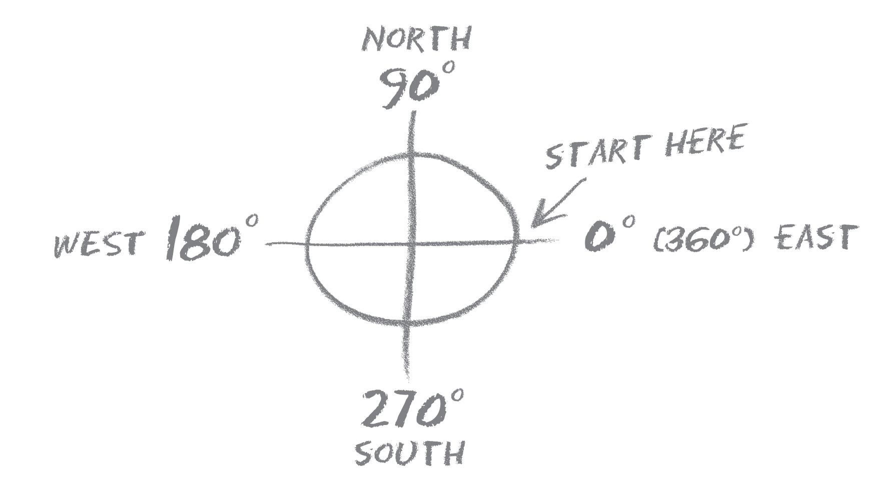


Resist the temptation to put values instead of percentages on the slices of a pie chart. If the actual values are essential to your business message, use another type of chart.

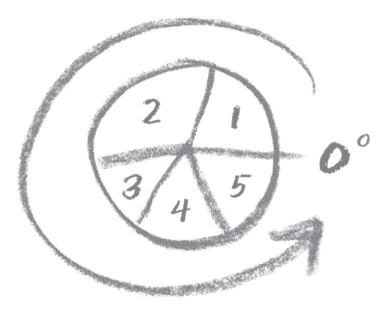
Pies, Clocks, and Compasses

A possible source of confusion about pie charts has to do with the *orientation,* or positioning, of the slices. There are two schools of thought on how to plot and read a pie chart. I'll call the two schools the "mathematicians" and the "clock watchers."

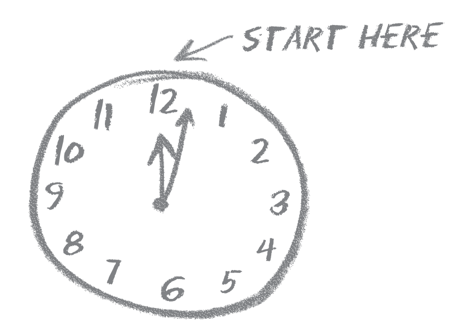
The mathematicians, following the rules of geometry, divide a circle (pie) into 360 degrees. The starting point of the circle is at 0 degrees (same as 360), which if you were looking at the points of a compass, would be due east:



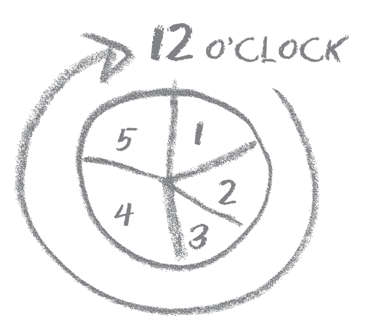
When they plot a series of percentages as pie slices, the mathematicians start at 0 degrees and proceed *counterclockwise*:



Some people, perhaps the mathematicians and their ilk, have been trained to read pie charts this way—counterclockwise. The clock watchers, on the other hand, think of a pie as a clock face:



The starting point is not at the side, but at the top—12 o'clock. Slices are plotted around the dial—you guessed it—in *clockwise* order. This is also the order in which the clock watchers read the data from the chart.



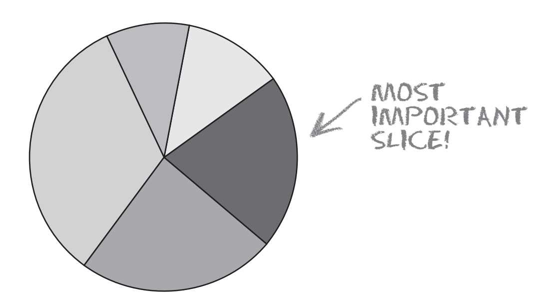


Despite the fact that more people know how to tell time than know geometry, some computer graphics programs use the mathematicians' method: *Start at 0 degrees and go counterclockwise.* Your audience, however, will most likely be populated with clock watchers: *Start at 12 o'clock and go clockwise.* Be aware of any automatic decisions the graphics program may be making for you, and adjust the results accordingly. Make your pies the way the important people in your life like them!

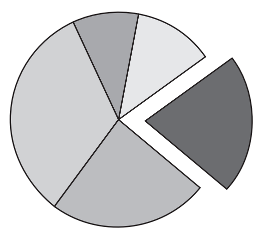
Which Slice Is Most Important?

Does the way a pie is built affect how the audience reads it? After all, don't people see the whole pie at a glance? They don't necessarily read the slices in any order. You might also suspect that the eye will be drawn to the biggest slice. The importance your audience gives to a particular slice, however, can depend on its *position* in the pie, regardless of its size.

Many people will assume that the most important slice is on the right side of the pie. To the mathematicians, that's 0 degrees. To the clock watchers, it's 3 o'clock:



The importance of this slice can be emphasized further by *exploding* it, or pulling it out from the rest of the pie:





Whether you are a clock watcher or a mathematician, the most important slice is always the one in the upper-right-hand position, *with this exception:* Regardless of its position in the pie, *any* piece that gets exploded becomes the most important piece!

A friend of mine made the following observation about the relative importance of pie slices:

> I was reading the *New York Times* this morning and discovered for myself how the 0 to 90 degrees position is where the eye seems to be drawn first. A pie chart showed the racial breakdown for people on welfare, and my eye was drawn to the group labeled "White" on the right-hand side, even though the percentage was about the same as the group labeled "Black." (The two groups combined are about 76 percent of the total.) It occurred to me that if the African-American group had been put instead in the top-right position, the chart might have emphasized a different point. My reaction—with the "White" group on the right-hand side—was: "That's interesting. The stereotype is that it's mostly minorities on welfare, but the percentage of white people on welfare is just as high."

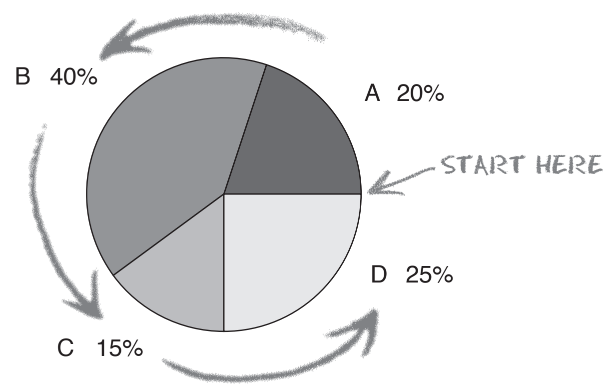
The *Times*, by the way, did not fall into the trap of showing numbers—they showed only the percentages.

What Determines the Order of Slices?

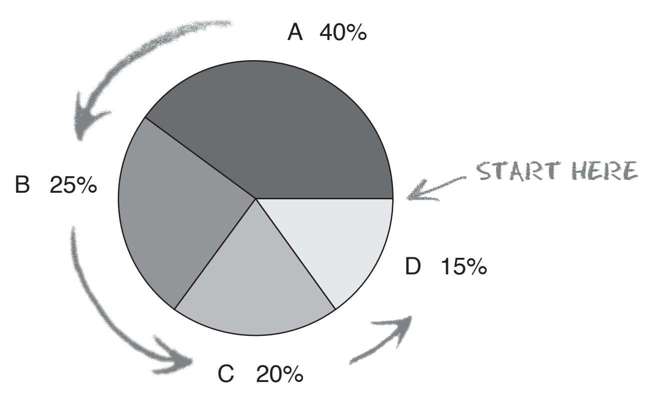
The order of pie slices normally follows the order of the values in the pie's series of data points. The order is simply the sequence of the data in list form, as the numbers might appear in a table or spreadsheet. For example, here is a series of percentages:

Slice	Percent
A	20
B	40
C	15
D	25

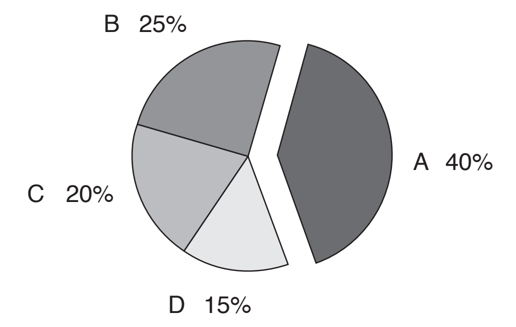
Using the counterclockwise mathematicians' method, this series will be plotted as:



Still proceeding counterclockwise, you might prefer to plot them in order of size:



Assuming that the 40 percent slice is most important, it can be emphasized by positioning it at 0 degrees—*rotating* the pie so that the right end of the pie is in the middle rather than at the beginning of the slice. Exploding the slice enhances the effect:

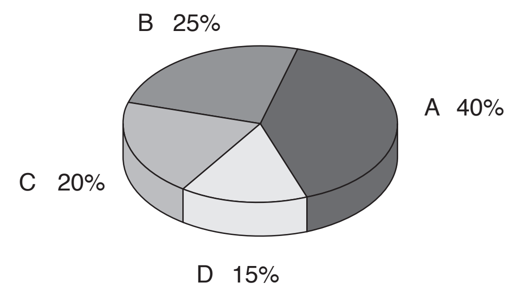




Of these examples, there is no "best" way to draw pie charts. You simply need to be aware of the hidden assumptions behind their construction. For example, if I were trying to downplay the amount my company was spending on vacations for the senior executives and play up the amount spent on charitable contributions, I could place the Charity slice in the right-hand position, exploded. Charity would then be emphasized, even if its percentage were smaller than the Vacations slice.

Which Slice Is Most Important? (Revised Version)

Now, the foregoing advice is all sound—until you start dealing with so-called *three-dimensional (3D)* pies. A 3D pie appears tilted on its edge so that it has thickness. The pie itself is no longer a circle—it's distorted, squashed into an ellipse:



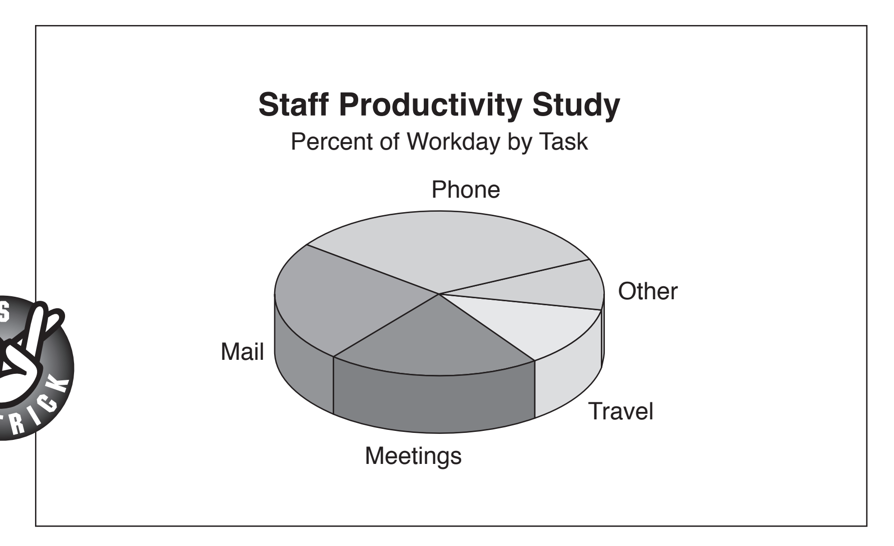


Figure 2.3 *This 3D pie really distorts the importance of the bottom slice. The thick edge makes the Meetings slice look much more substantial than the Phone slice, even though its percentage is smaller.*

In a 3D pie, people will usually think that the *bottom slice* is the most important. That's because the dimensional effect distorts the apparent size of the slice by literally giving it an edge. People will perceive this slice as bigger than it would appear in a circular, two-dimensional pie.



Here's one of the liar's favorite tricks: Put the slice you want to emphasize at the bottom of a 3D pie. Give the pie some real thickness and tilt it way back. And, to compound the crime, omit all data labels. The effect is awesome, as you can see in Figure 2.3.

Not Using Pies to Show Sales Results

Perhaps the most common—and the most offensive—misuse of pie charts is for showing sales results (see Figure 2.4). This example violates my cherished rule about pies: The slices are labeled as dollar values rather than as percentages. The audience will be strongly tempted to mentally add the slice values to find out the total sales dollars, especially if the amount has not been shown previously in the presentation.

A correct use of a pie would be to compare market share percentages among competitors, as shown in Figure 2.5. This is the same pie as

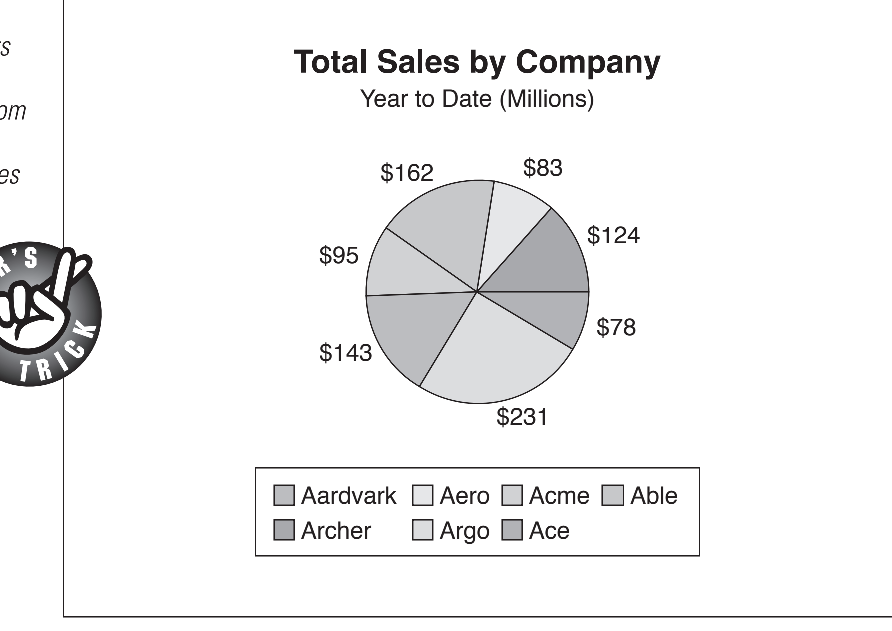


Figure 2.4 *Including dollar amounts in pies that show sales results is usually a bad idea because it distracts from the real story—the* proportions *of the slices to the whole.*

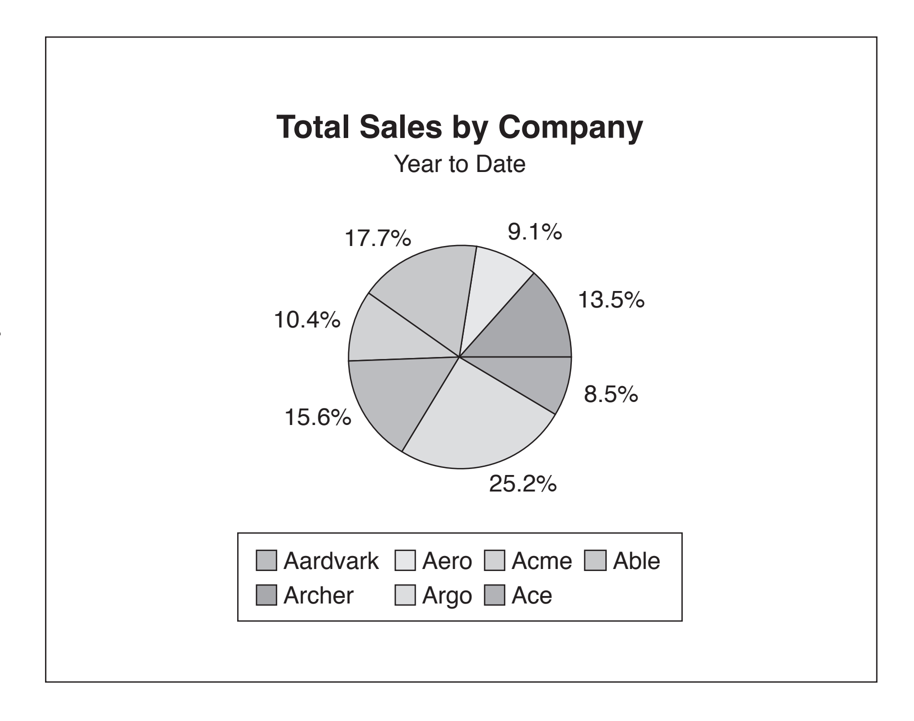


Figure 2.5 *Here is the chart from Figure 2.4, but with percentage labels instead of dollar amounts. When pie slices are labeled with percentages, the audience will focus correctly on the relative importance of the slices rather than on their explicit values.*

Figure 2.4, but with percentage labels. The audience is not tempted to ask about the total sales dollars—the volume is not the issue. Instead, attention is focused on the percentages of the total market.

Don't Tell the Two-Pie Story

Another trap opens wide for you when you proudly carry your pies into the world of sales. Whenever you are dealing with sales results, you can't expect your audience to remain focused on the present. They also want to know about the *past* and the *future!* Today's sales results are always measured against yesterday's—or last month's or last year's. And, before you finish your presentation, the audience will naturally be interested in your projections about tomorrow's sales.

Sales results always get this yesterday-today-and-tomorrow perspective. That's because sales dollars represent the ongoing income of the business—its very life. *Everyone* in a business is interested in actual sales dollars because fluctuations in income affect everyone's financial well-being.

For these reasons, one pie used to show sales dollars can quickly become *two*—LAST YEAR and THIS YEAR—on the same chart, as shown in Figure 2.6. Bolder presenters will venture to try *three*—including another pie with estimates for NEXT YEAR.

This two-pie comparison is confusing because it uses sales dollars—actual values—as slice labels. It breaks the rule about using pies to show only percentages. As with the single-pie example, this chart would make more sense if it were intended to show market shares only—dropping the dollars in favor of percentages.

Nevertheless, heedless tellers of two-pie stories will stop at nothing! Making both pies the *same size* looks wrong, they think. After all, this year's sales are greater than last year's—so, they reason, this year's pie should be bigger! Compounding the error of showing the sales dollars instead of the market-share percentages, they adjust the sizes of the pies, as shown in Figure 2.7.

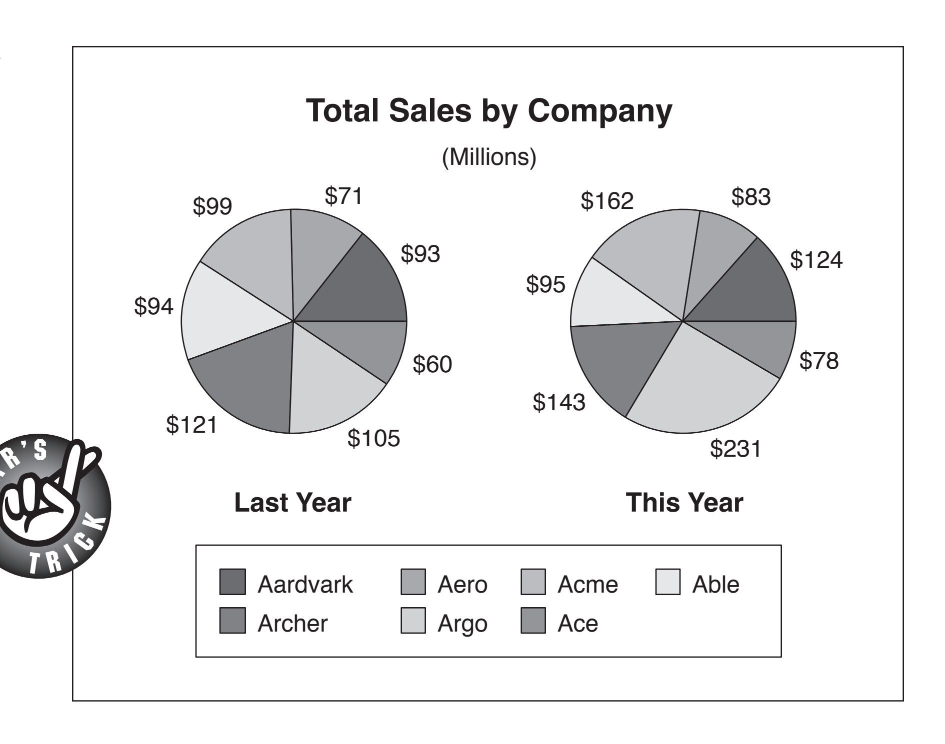


Figure 2.6 *The Two-Pie Story is told by information gluttons who can't resist the temptation to have that second pie!*

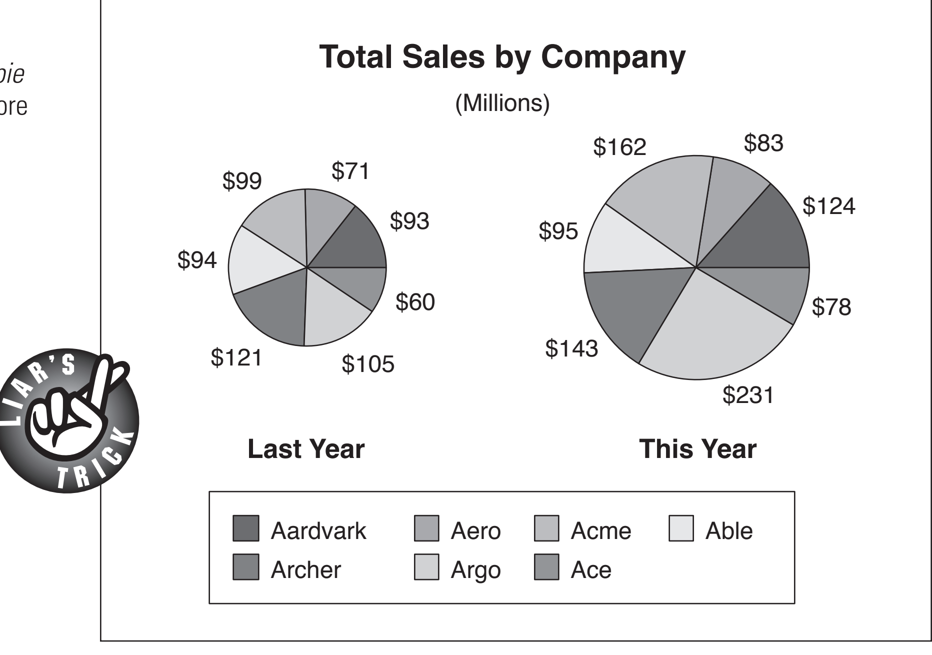


Figure 2.7 *Using different-size pies just makes a two-pie chart that much* more *confusing.*

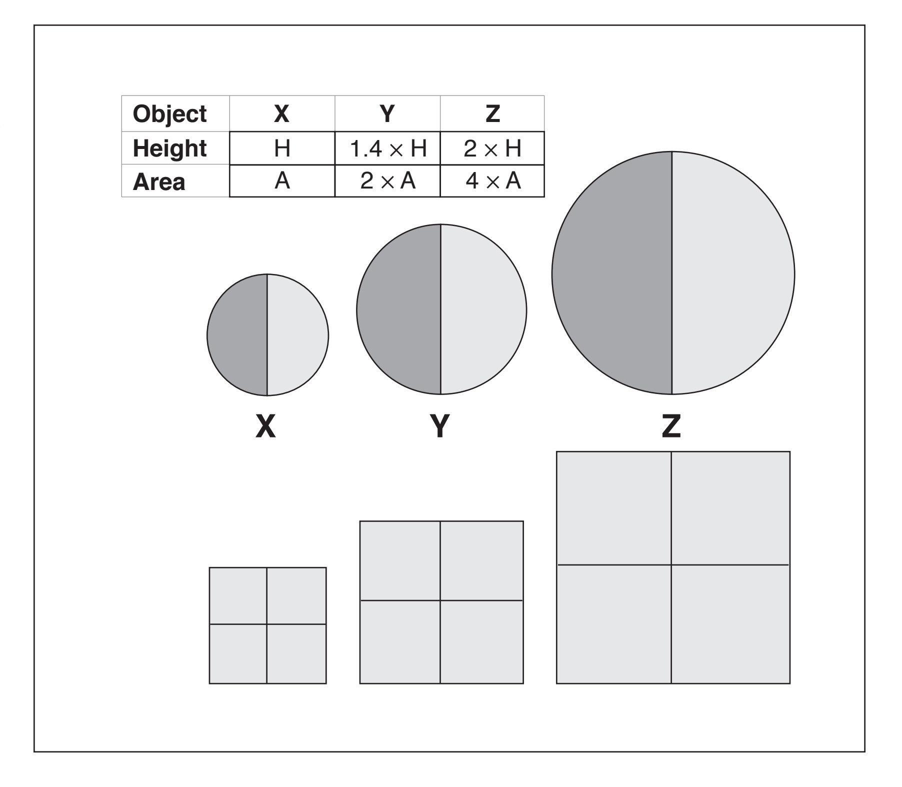


Figure 2.8 *An accurate comparison of the sizes of objects—whether circles or squares—is by area, not by height.*

$$$$INVESTOR'S TIP$$$$

Consider that doubling one dimension of any 2D shape while preserving the aspect ratio effectively squares its area. So it's just as misleading for area charts or symbols as it is for circles and pies.

Some people will be perfectly happy with the chart in Figure 2.7. It's a feast for fools, though.

Here's why: If this year's sales are 1.4 times greater than last year's, the chart designer's inclination will be to treat them like bars and make the height (the *diameter)* of the This Year pie 1.4 times that of the Last Year pie, as shown in Figure 2.7.

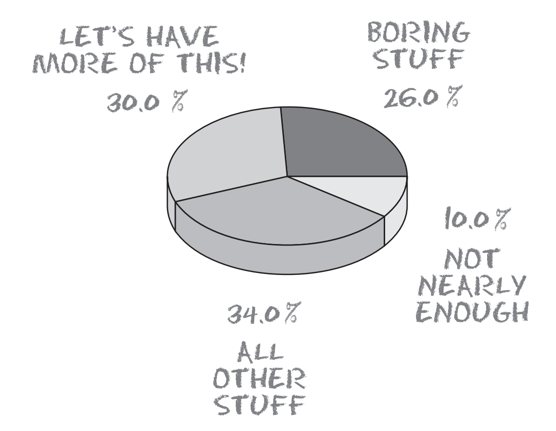
However, an accurate comparison of the sizes of objects—whether circles or squares—is by *area,* not by height. In Figure 2.8, the area of the pie on the right is actually *twice* that of the pie on the left. The correct sizing of the second pie—which is a smaller area—is shown in the center. The two-pie chart is misleading because the second pie is larger than it should be, inflating the sales dollars it represents. The true relationship of the areas is easier to see when they are drawn as squares, as shown in the lower part of the figure.

Further Secrets of Pie Thieves

Liars and pie thieves have a few other tricks that they can do with Mom's apple pie:

Adjust Overall Pie Size

Even if you are showing just one pie, you can distort the importance of the whole issue by playing with the size of the pie. Liars who want to diminish whatever the pie represents will often clutter the chart with lots of labels, forcing the pie to be smaller to fit on the page or screen.



Abuse the All Others Category

In a pie chart, you are forced to show the whole thing, whatever the thing under discussion is: The whole pie always represents 100 percent. Liars sometimes find it inconvenient that all pieces of the pie must be accounted for. They typically get around this by labeling a mystery slice All Others.

MY STUFF!
80%

ALL OTHERS
20%

Watch out when you see All Others, especially if the slice is more than 10 or 15 percent of the whole pie. The presenter just might be hiding a big one in there, rather like burying your vacation in the Business Travel category of your expense report.

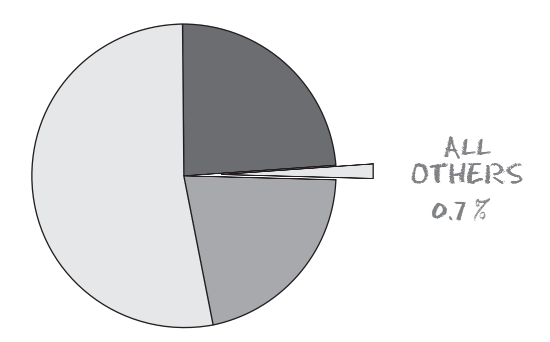
There can be legitimate reasons to have an All Others category. For example, if you are showing market share, there could be a dozen competitors that share just 10 percent of the market—and, for the moment, none of them is important to you.

Don't Show the "Don't Knows"

If the All Others slice complicates your story, simply exclude it from the pie! The effect will be to increase the sizes of the other slices. Their percentages will be increased because you are reducing the amount of the unseen total. Most audiences won't catch on that you're not telling them the *whole* story!

Fudge Small Slices

There is one situation in which I *recommend* that you cheat in preparing a pie chart. If you must include a very small slice—equivalent to 1 percent or less—I recommend plotting it at about 1.5 percent (about 5 degrees of the circle) so that the thin slice can be seen. Otherwise, a thin slice looks too much like the *callout lines* that some designers use to connect labels to slices. You should include a helpful label nearby that clearly shows the true percentage. Exploding the slice, as shown here, may help you make your point.



Exotic Pies for the Info Gourmet

Two variations of the pie chart can be especially tempting on special occasions: the pie-column chart and the donut chart.

The Pie-Column Chart

This hybrid chart format combines an exploded pie with a *column*. The column is a type of stacked bar chart, its segments stacked on top of one another. The column presents an itemization of the exploded slice, as shown in Figure 2.9. This example is a solution to the All Others problem described above: Show the market shares of all those tiny competitors in a column.

You might also use a column to itemize and focus in on a breakdown of the biggest or most important pie slice.

Some people use a second pie instead of a column, but I think the columnar design is less confusing to the audience. It will be obvious that the column is not a second version of the pie.

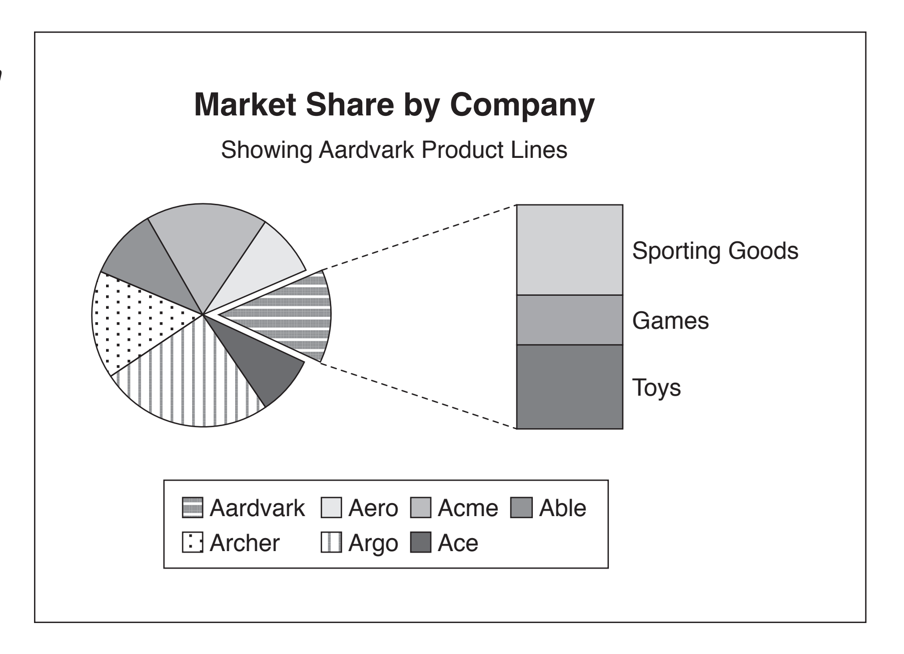


Figure 2.9 *In a pie-column chart, the column itemizes the exploded slice.*

The Donut Chart

Some clever designer decided to hollow out a bunch of pies and show them stacked inside each other—like rings on an archery target. The prize-winning name for this concoction is the *donut chart,* as shown in Figure 2.10.

Now, I'm not sure what use a donut chart might have for you, except perhaps to perplex your audience. (Liars take note!) It's just that pies are so simple and basic—easy to digest, you might say—why not use a series of pies on different charts rather then trying to cram them all around the same little donut hole?

But, if your audience likes these donuts, go ahead and feed them what they want. Your presumably legitimate purpose might be to compare the percentages, or ratios, of several different entities. For example, look at Figure 2.10. If you were showing market shares, each ring on the donut might represent comparative market penetration in a different product line. Just remember to give your audience plenty of time to study this one! And, as with pies, avoid labeling donut slices with explicit values.

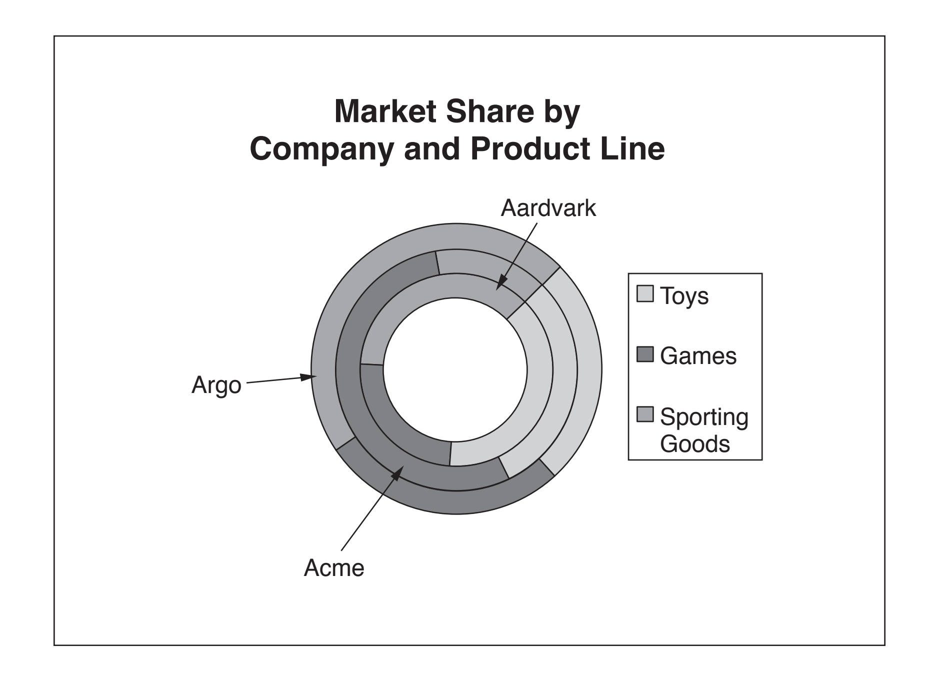


Figure 2.10 *A donut chart is a set of concentric, hollowed-out pies.*

Honest Alternatives to Rotten Pies

Remember, pies are for percentages or ratios. If you find yourself putting actual numbers (such as dollars) on a pie, you might be able to clean up your act by using bars or areas instead. Here are some of the alternatives:

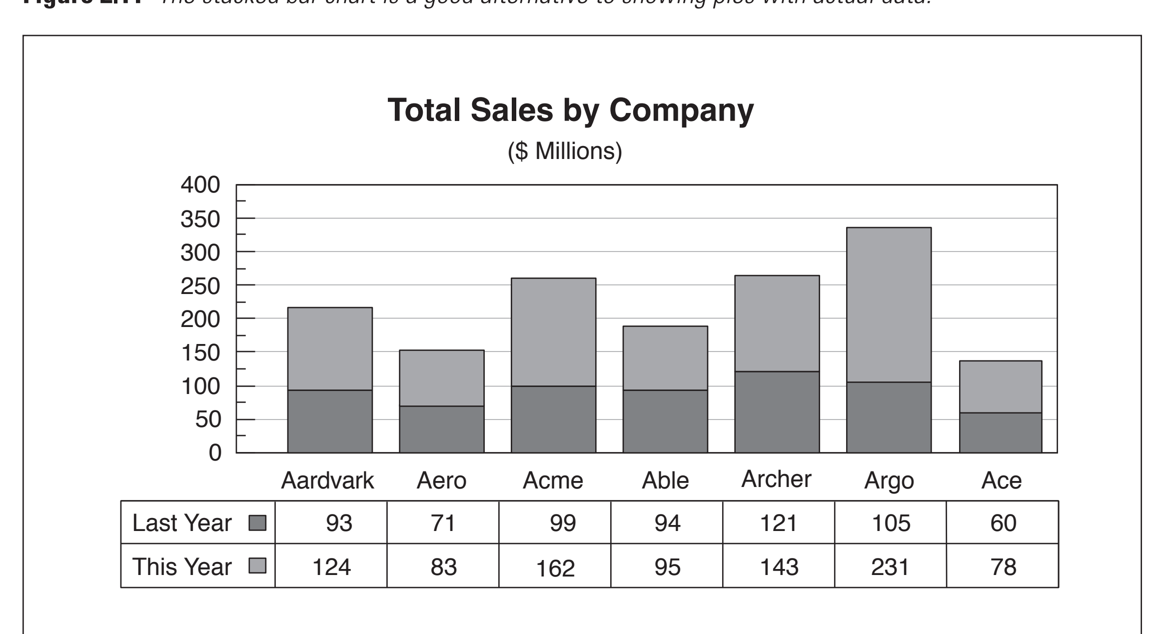
Stacked Bars

As shown in Figure 2.11, this is the best way to show relative sizes of different segments as well as their actual amounts. (There's more about these stacked things in Chapter 4.)

100 Percent Bars

In this format, all stacked bars are the same height, each representing 100 percent (see Figure 2.12). Think of them as pie dough rolled

Figure 2.11 *The stacked bar chart is a good alternative to showing pies with actual data.*



	Aardvark	Aero	Acme	Able	Archer	Argo	Ace
Last Year	93	71	99	94	121	105	60
This Year	124	83	162	95	143	231	78

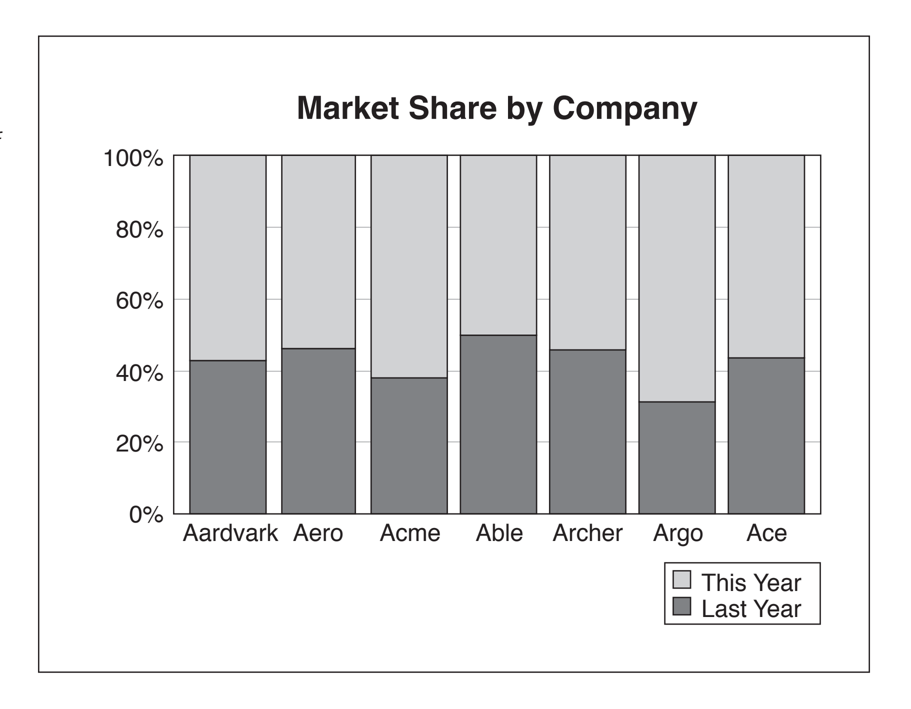


Figure 2.12 *The 100 percent bar chart can be used instead of a donut chart or instead of multiple pies.*

into same-sized bars. As with donut charts, you can compare the equivalent of several pies in one chart. As with pies, the labels should be percentages, not actual values.

Stacked Areas

Like stacked bars, this format can be used to show both relative sizes and actual values (see Figure 2.13). Deciding the relative positions of the areas can be confusing to an audience, however. Are the areas stacked vertically—like bricks and mortar—or are they layered on top of one another—like drawings on separate transparent sheets, each new area starting on the same horizontal line at the bottom of the chart? If the areas are truly stacked, only the bottom data series will have a flat baseline. Each new area will take the top of the preceding one as its baseline, and any fluctuations in that line will distort the highs and lows in the series above. (These distortions can be minimized by putting the "least bumpy" areas on the bottom.)

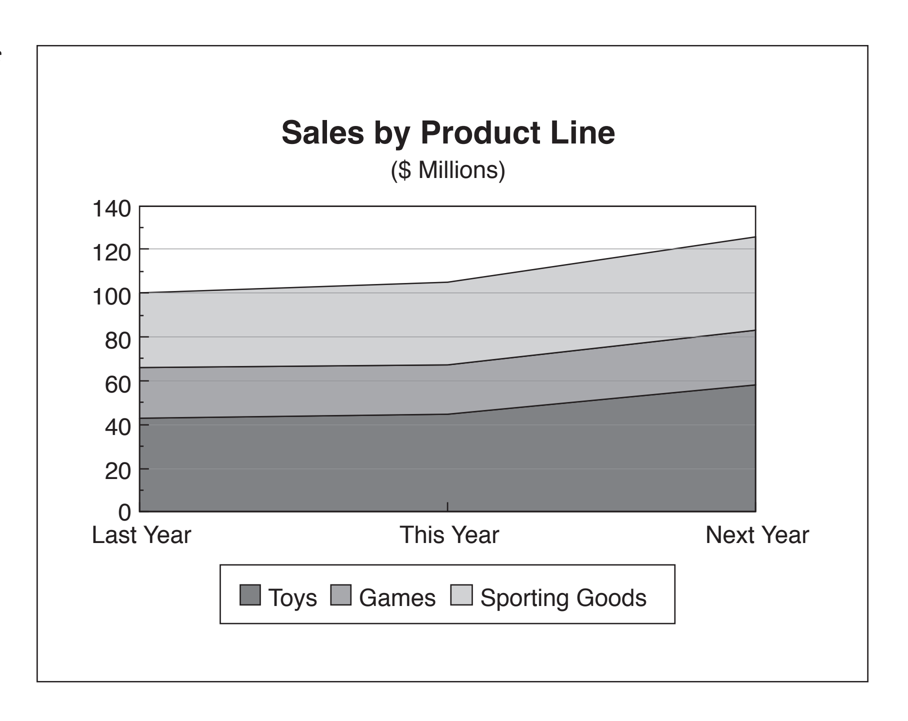


Figure 2.13 *Area charts are alternatives for pies, but where's the baseline of each series? Is it the bottom of the chart, or is it the top line of the preceding series?*



There's much more about bars and areas—including the stacked kind—in Chapter 4.

Okay, What Are Pies For?

The pop quiz for this chapter has just one question. And you know the answer:

Pies are for percentages!

If by now you're sick of my harping on this point, then I guess I've done my job!

3

Orientation

Or, What Goes Up Must Go North

CHARTS are pictures. In this chapter, I encourage you to think about them that way—as collections of forms rather than diagrams of data. In particular, I want you to think about *orientation*—the position and direction of graphic shapes in relation to a page or screen. Artists call it *composition,* and it can have powerful effects on an audience, regardless of the data content of your charts.

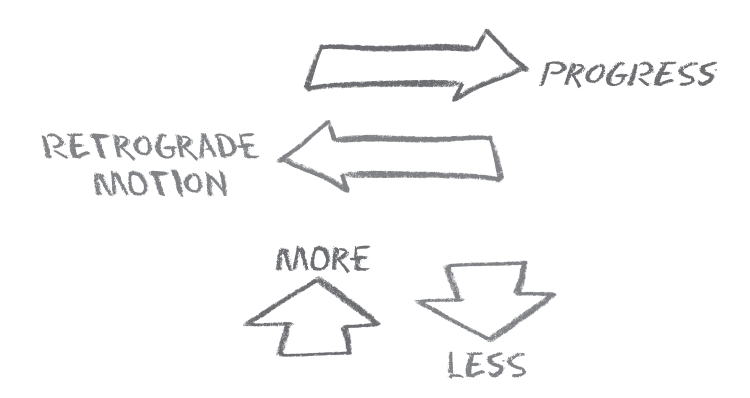
When you set out to make a chart, you should be aware of some basic assumptions that are shared between the presenter and the audience. These (usually subconscious) assumptions include the meaning and importance of directions in space. For any kind of graphic, designers must make decisions—conscious or not—about how the shapes that compose the image are oriented. Viewers get very different impressions depending on whether shapes appear to go left, right, up, or down.

Where Does the Time Go?

People in Western cultures read from left to right. For these readers, rightward motion—the way the eye scans a page—is associated with the passage of time, and thus with positive movement, or even the idea of *progress.*

$$$$INVESTOR'S TIP$$$$

The notions of time flowing rightward and value increasing upward is inherent in all stock market charts. There is no other expected scheme.



Left Is Just, Well, Gauche

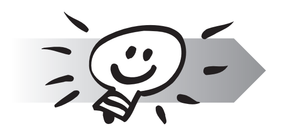
Conversely, movement from right to left is considered backward and negative—*bad!* From these same notions came the root of the English word *sinister*, which means both *left* and *unlucky*.

As you will see, charts that show the progress of time rely on the notion that it flows from left to right.

Up Is More, Which Eventually Comes Down, Becoming Less

In most cultures, movement upwards is associated with increase or gain, downward with decrease or loss. Where charts are concerned, up-down orientation has another effect: People who are used to getting their information from reading text will often begin at the top of a screen and scan downward, even if they are looking at a graph rather than reading words on a page. This downward movement of the eye is not associated with negativity—you don't feel worse about the words at the bottom of the page. It's simply that readers are trained to look from top to bottom, so they will do this just as surely with a chart as with text.

Charts that show fluctuations in amounts rely on the notion that *up* means *gain* and *down* means *loss.*



Concepts of orientation aren't necessarily the same when you are looking at photographs. Scientific studies have shown that eye movements involved in looking at photographs or live scenes are very rapid and complex—especially when the viewer is studying another human face. More about faces at the conclusion of this chapter.

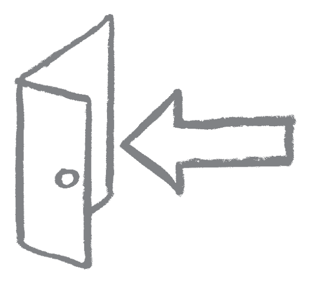
Be Glad Your Audience Is Prejudiced

Much effort in the graphic design field has been devoted to finding more universal means of expression to minimize the effects of biases, which are primarily cultural. Cultural biases are things you learn—rules of a particular society—not instincts you're born with.

International signs constructed of simple, generic symbols are examples of designs that attempt to minimize some cultural biases. The designs aim for simplicity, reducing the details in symbols to the bare essentials—removing from human figures, for example, such details as skin color and apparel, as well as eliminating explanatory text wherever possible:



Because international signs must be cross-cultural, they can't take advantage of orientation. Nevertheless, the effects of orientation can't always be avoided. These biases exist apart from the literal interpretation of a sign. For example, this sign is cross-cultural—even people with different cultural backgrounds will interpret this sign to mean something like: "This way to the exit."



Beyond this literal interpretation, viewers of a sign may respond emotionally to it, according to the orientation of its symbols—in this case, the direction the human figure is walking and the position of the arrow.

It's What You Don't Say

Which of these signs might make a Western traveler feel more "positive" about a choice of routes?



Never mind that there is only one spot to place the sign in an airport terminal and the arrow *must* go the other way. Certainly, you can't go pointing people in the wrong direction just to make them feel better. And no one would ignore a helpful left-pointing sign because of such a vague feeling. But, be aware that people carry those biases with them, no matter which way they decide to turn. You might not have any choice about which design to use in the airport situation, but you could give careful consideration to how you would use the same sign inside a store to direct shoppers to a merchandise display.

When your audience shares a set of biases, a chart will be more effective if its design plays to those assumptions deliberately. For example, you can probably assume that Westerners will interpret rightward motion as progress. *Use that prejudice to your advantage!* Don't try to neutralize their biases, which probably isn't possible anyway. Simply make your designs consistent with what they expect.

It might seem somehow unethical to reinforce these prejudices, but realize that you can't escape them. Graphic conventions are as much a part of language as speech. Your choice of the language itself depends upon many conventions, including these notions of direction. Different sets of assumptions *and visual biases* will apply, for example, to native speakers of Hebrew, Arabic, and Farsi (who read right to left) or Chinese (who sometimes read top to bottom).

In short, your charts aren't international signs. You can't make them cross-cultural, and you shouldn't try. They will have more impact if you are aware of cultural biases and play shamelessly to them!

Will You Get Rave Reviews in Tel Aviv?

What should these differences in culture mean to you, in practical terms? For one thing, don't assume that if a presentation made a big hit in the New York office, it will make the same favorable impression at an international conference in Tel Aviv.

Wait, you say, isn't charting an international language of its own? Like mathematics?

Well, chart formats generally transcend language.

But, like the interpretation of international signs, the concern remains how the audience will *feel* about a chart, beyond the literal interpretation of its data.

Why do you need to know these things? Well, if you make a chart that shows rising sales, you want its composition to convey the reinforcing feelings of positive motion and gain. But if you ignore the effects of orientation, your chart might instead convey the opposite feelings of retrograde motion and loss! Ambitious liars can learn lessons here, as well. They can make bad news look not so bad (to an unsuspecting audience).

XY Means Quantity Versus Time (Usually)

A set of charting conventions that play directly to Western cultural bias is the *xy* chart, perhaps the most common format used in business presentations (see Figure 3.1). *The* xy *chart is a powerful device precisely because it is rooted in cultural biases about the meanings of left-right, up-down.* The how-to of charting in *x* and *y* is covered in Chapter 4. My focus here is on the messages that might be hidden in these charts.

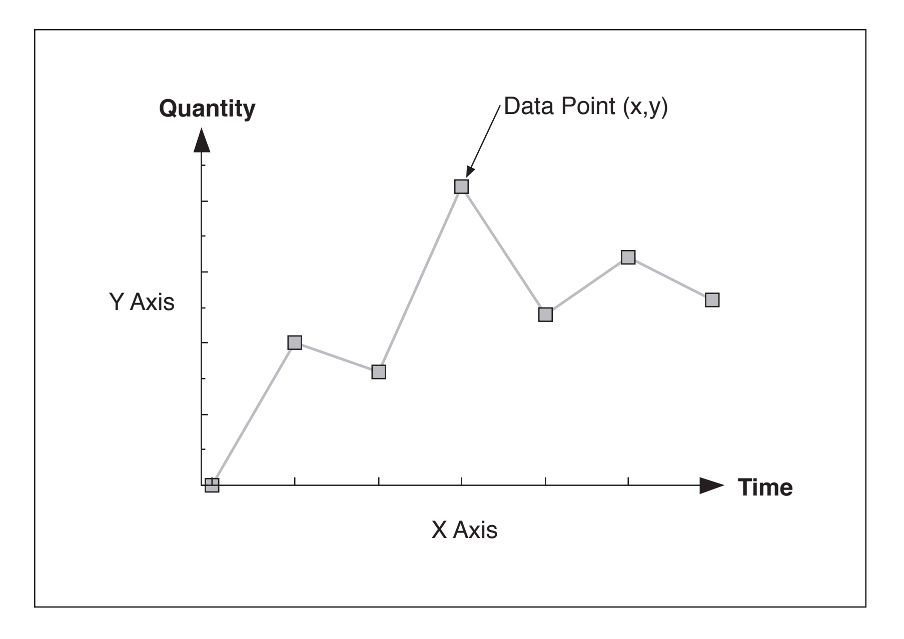


Figure 3.1 *To most Western audiences, an* xy *chart shows a relationship between a variable quantity and the progress of time.*

Bar, line, area, and scatter charts are all based on *xy* plotting conventions. They all show *quantity* in relation to *time.* Quantity goes up and down on the vertical axis—the *y axis*. Time progresses from left to right along the horizontal axis—the *x axis*. People who share the assumptions that *up* means gain and *right* means progress (or just "in the future") can learn much about a trend line even if the graph has no precise detail in the form of scales or labels.

For example, the crude graphic in Figure 3.2 conveys a simple message: "The U.S. national debt has been growing especially rapidly in recent years, and threatens to get out of control." (The chart certainly implies this, although historically the situation is a bit more complicated. You'll see what I mean shortly.)

Look again at the text of the message. That's a lot of words, although far short of the thousand the picture is supposedly worth. Still, the text of the interpretation is a surprisingly specific conclusion from little more than a squiggly line and a couple of symbols!

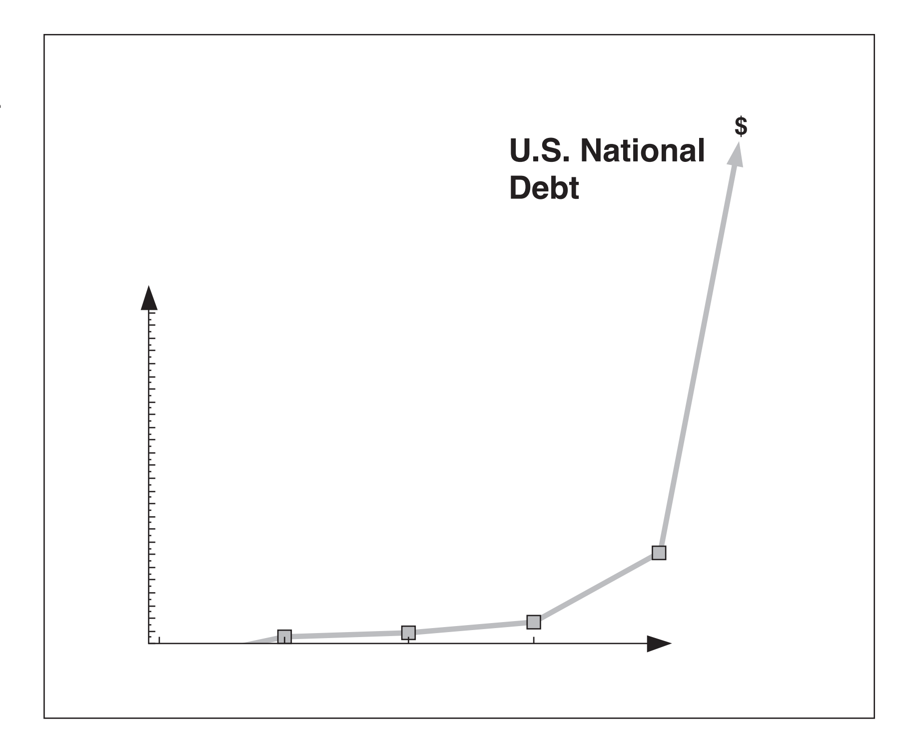


Figure 3.2 *This crude chart, which has no numeric scales or labels to help you interpret it, has a powerful implied message.*

It's Obvious Because It's ... *Um*... Obvious

The fact that this example seems so obvious is, in itself, proof of the unspoken assumptions it uses so effectively. Stop a moment to consider what you *don't* see in that graph. How do you know it's an *xy* graph? Furthermore, how do you know it's a *vertical* xy (quantity goes up-down not left-right)? Is there anything *in the graph itself* that tells you the events plotted at the left end of the line are further in the past than events at the right end?

Now, look at the chart in Figure 3.3. What is the message *here?*

The answer is: The same as the first chart (Figure 3.2)! They are both plotted from the same set of data.

In the second example, the scales have been rotated so that the plotting orientation is horizontal rather then vertical. Nevertheless, this

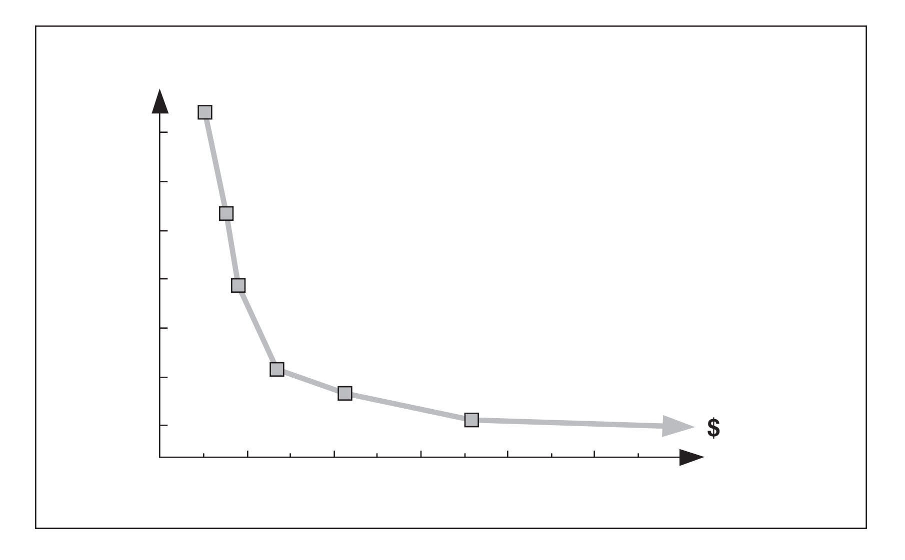


Figure 3.3 *This chart plots the same set of data shown in Figure 3.2, but because of its unconventional orientation, the intended message is not as obvious.*

odd-looking version is a perfectly valid chart. It's just not the format you're used to seeing *for this type of message*—quantity versus time. Perhaps it will become clearer if the odd-looking chart is shown with numeric scales and descriptive labels, as in Figure 3.4.

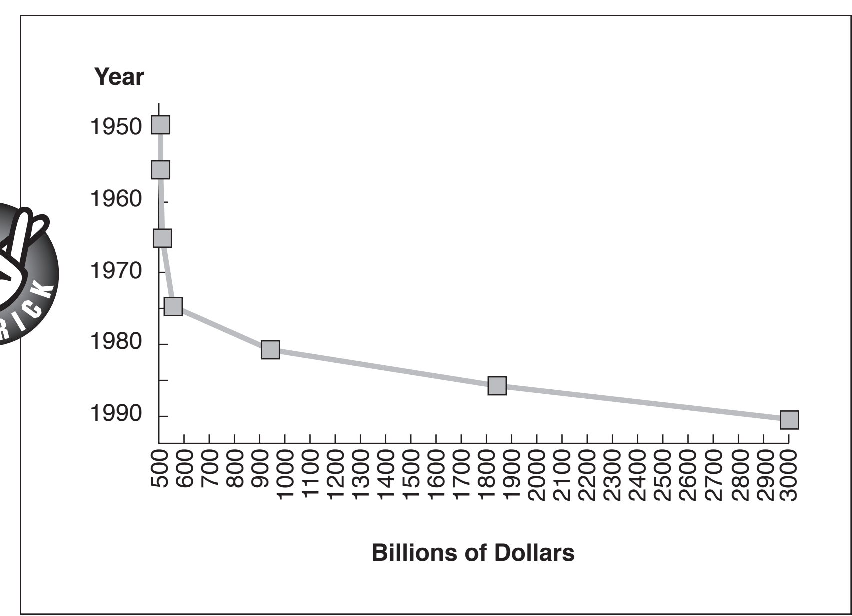
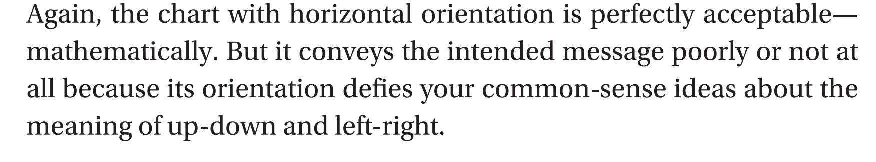


Figure 3.4 *Adding scales and descriptive labels makes it clear that this chart, despite its odd format, tells a much more accurate story: In the second half of the 20th century, the U.S. National Debt skyrocketed, with most of the increase occurring in the 1980s. In more recent years, not shown on this chart, the debt went down abruptly, then shot right back up again.*

LIAR'S TRICK



Again, the chart with horizontal orientation is perfectly acceptable—mathematically. But it conveys the intended message poorly or not at all because its orientation defies your common-sense ideas about the meaning of up-down and left-right.



Many chart types do not adhere to the time-magnitude relationship just described. A particularly weird format is the *radar chart,* which is so intriguing I cover it separately in Chapter 5.

As a presenter, you depart from convention at some risk. It generally would not be appropriate, for example, to use a horizontal bar chart to represent sales volumes, as in Figure 3.5.

$$$$INVESTOR'S TIP$$$$

In the Microsoft jargon of PowerPoint and Excel, a horizontal bar chart is the default bar format. To sharpen the distinction, a vertical bar chart is called a *column chart.*

Your intention might be to introduce variety to the presentation, but the result will be added confusion. Horizontal bars are an excellent way to show time spans and durations, however. Figure 3.6, which shows how many hours per week staff spend on various tasks, is a good example of a chart that uses horizontal orientation effectively.

Figure 3.5 *Using a horizontal orientation to show sales volumes fails to capitalize on the intuitive notion that* up *means* more *in a vertical chart. However, using horizontal bars* can *give the impression of progress toward a goal.*

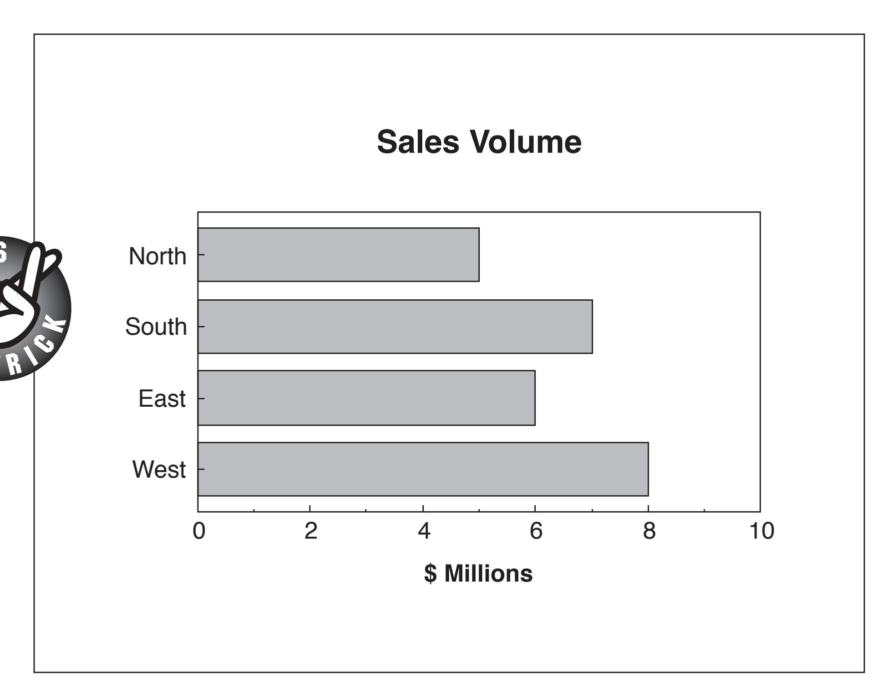
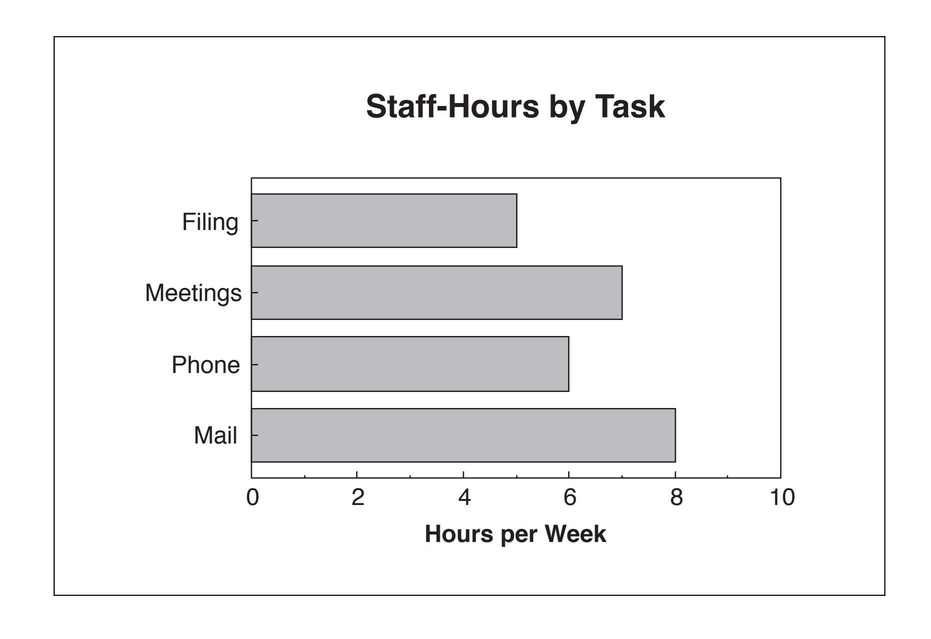


Figure 3.6 *Because left-to-right motion generally implies the passage of time, a chart with horizontal orientation is ideal for plotting time spans.*



The *Gantt chart,* or schedule chart, is a specific format used in project management for showing task durations and timing relationships as horizontal bars (see Figure 3.7).

Figure 3.7 *The* Gantt chart *is a specialized horizontal charting format used to show the durations of individual tasks within a larger project. It also capitalizes on the intuitive notion of Western readers that time flows from left to right.*

Project Schedule

ID	Task Name	Start	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
1	Plan	Mar 15													
2	Develop	Mar 22													
3	Test	Apr 5													
4	Implement	Apr 12													



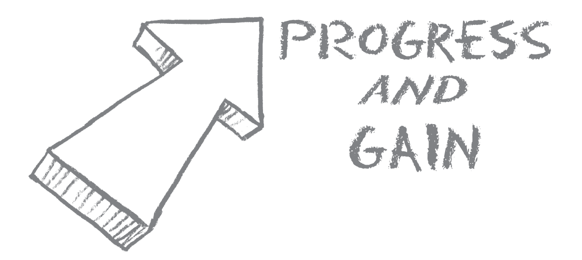
Specialized software products have been developed for project management which typically can generate Gantt charts from a scheduling process called *time analysis.* The chart in the figure was generated by Microsoft Project. Even though Gantt charts might not be specifically supported by some graphics programs, there is usually something like a *stacked horizontal bar* chart which can be adapted for this purpose. There's more about Gantt charts in the next chapter.

Tempted to Try a New Slant?

So much for right angles. What about slanting lines? What impact do they have?

Any direction on a page or screen can be described in terms of left-right (horizontal position) and up-down (vertical position). *The impact of a slanting line is simply the net effect of its left-right and up-down components.*

For example, you know by now the hidden message in a right-pointing arrow, and you know what an upward arrow can mean. A straight, right-pointing arrow has no vertical component, no magnitude. Therefore, it implies progress in time only. An upward-pointing arrow shows an increase in magnitude, but because it has no horizontal component, or no time span, it implies a static situation. A third design uses an arrow pointing upward and to the right to convey a sense of *both* forward motion and growth, or gain:



This combination of both positive biases makes the slanting arrow by far the strongest visual statement of the three designs. The two positive components *up* and *right* combine to make a doubly positive

impression. Likewise, combining two negative components—*down* and *left*—results in a forcefully negative design:



Notions of North, South, East, and West

Linguistic and graphic conventions about left-right and up-down carry over to concepts about geography. These concepts should also be considered in your graphic designs—even in charts that have nothing to do with maps!

From early childhood, North Americans have been taught to visualize the earth in terms of a Mercator projection map with the continental United States at or near the center (Figure 3.8).

Since childhood, Westerners are taught their geography with reference to Mercator. So I learned that North is up, South is down, East is right, and West is left. But there's no objective reason to view the world that way: we're on a globe floating in space, where *up* has no objective meaning at all!

Figure 3.9 shows McArthur's Universal Corrective Map of the World, an upside down (or "down under") map drawn by some rebellious souls who apparently think Australia is the center of everything! Try thinking about how your assumptions about up-down-left-right would be different had you been taught this map in elementary school.

With the conventional biases of Mercator, travelers planning a trip from the United States to Europe often have a mental model of flying in an easterly, even rightward, direction. Some people are surprised to learn

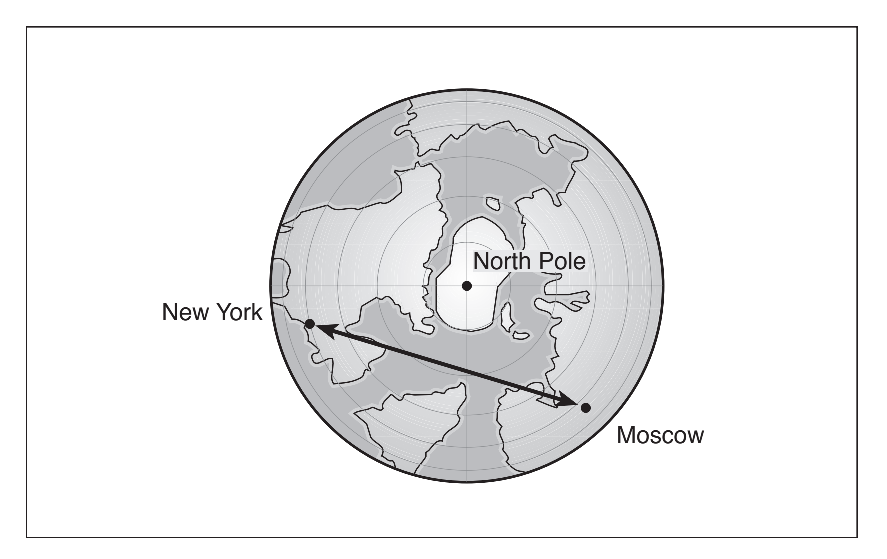


Figure 3.8 *The Mercator map of the world is so familiar that it affects many of our everyday notions about the meaning of left, right, up, and down.*



Figure 3.9 *McArthur's Universal Corrective Map of the World, drawn by self-centered Aussies, shows me as a North American how self-centered my own Mercator-inspired notions are.*

Figure 3.10 *A polar projection map shows that the air route between New York and Moscow is more of an up-down trip than a left-right one—all terms that have no objective meaning in terms of a globe.*

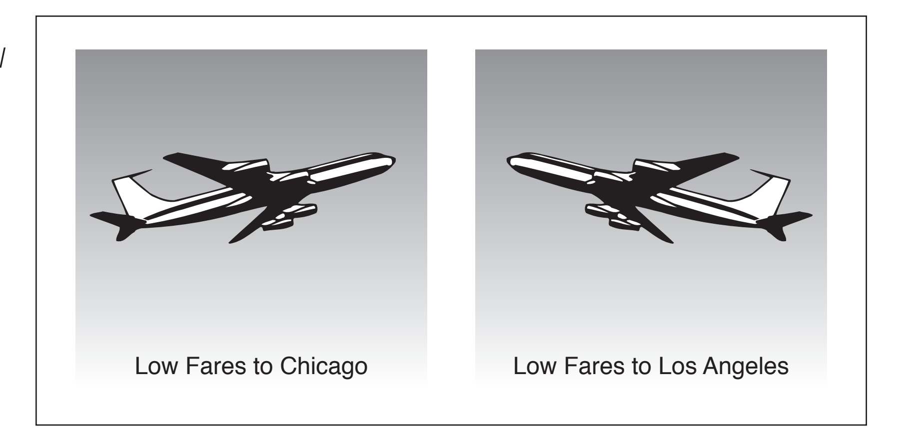


that the flight plan from most U.S. cities would follow a Great Circle route that goes "up," over the North Pole, as shown in Figure 3.10.

This cultural convention can also be seen in television advertising for air travel (see Figure 3.11) . In California, an airline that advertises flights between Los Angeles and Chicago shows an aircraft flying from left to right, with its nose pointed toward the upper right of the screen. Viewers in Los Angeles, who are thinking primarily about their departure rather than their return, perceive this screen direction as meaning *northeast*. It's fascinating to view similar spots that are prepared for Chicago audiences: As you might expect, the plane is pointed in the opposite direction—to indicate westward flight. In such cases, the art director is designing the scene to meet viewers' expectations.

The orientation of objects on a screen can even have visceral effects on an audience. Map conventions are so internalized that they influence involuntary eye and body movements. When you read a map, you have been taught to turn it so that north is at the top. Some of us will also go to the extreme of turning our bodies to face north while reading the map.

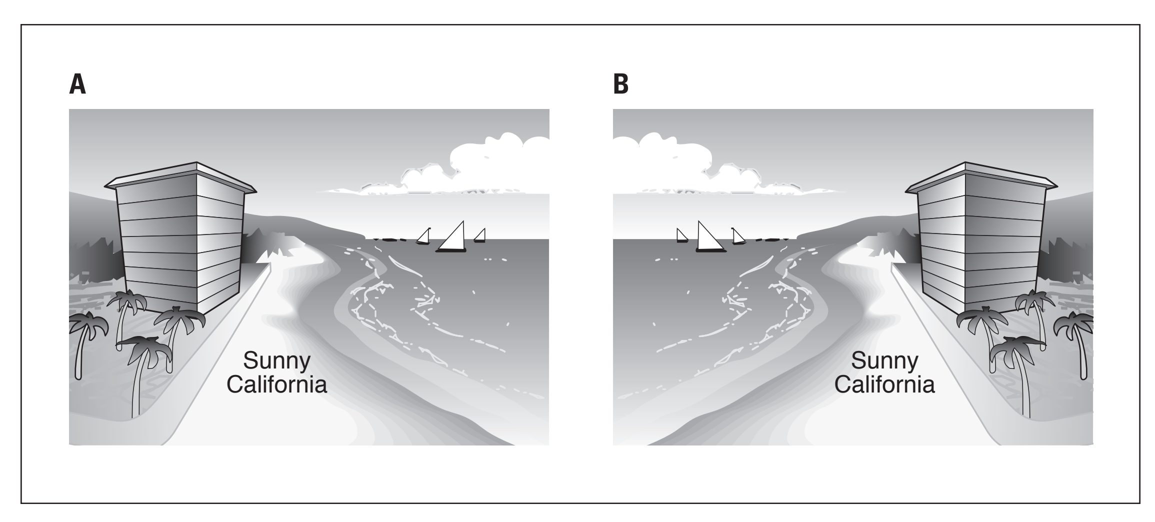
Figure 3.11 *In television, movies, and travel advertising, a journey to the right usually means* eastward, *left means* westward.



What's Wrong with This Picture?

Photographs or drawings that involve geography can also play to these prejudices about screen direction. Again, because of our dependence on Mercator, people in North America think of the Atlantic Ocean as being on the right and the Pacific on the left. If this seems silly to you, look at the sketches in Figure 3.12.

Figure 3.12 *Many people will think that picture A is wrong, although they might not be able to tell you why. Because of Mercator, they have the notion that the Pacific Ocean should be on the left, as it is shown in picture B.*



Lessons for Chart Makers

Again, distinctions of geography and direction are purely arbitrary. The real information content of airline commercials or the beach scene in the preceding examples is not affected by screen orientation. However, to the extent that people have deep-seated expectations, you should use them to your advantage.

And don't ignore them! If these prejudices aren't working for you, they are working against you.

For example, the chart in Figure 3.13 shows sales by region. The audience will be uncomfortable, and therefore needlessly distracted, because the West is not on the far left, Central in the middle, and East on the right of the *x* axis.

Don't get me wrong. It's not important to list the sales regions in any particular order for the sake of the honesty or integrity of the graph. The composition is just less effective because it ignores the prejudices of the audience. The chart might make them uncomfortable, even if the effect is only subconscious. And, because they are not focusing completely

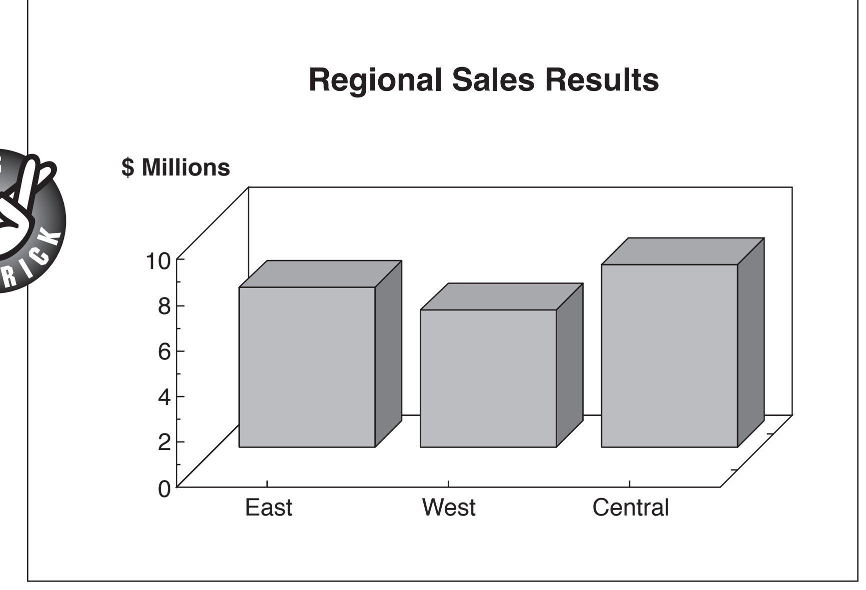


Figure 3.13 *To meet viewers' expectations about direction and geography, the maker of this chart could have put the West on the left, Central in the middle, and East on the right. (Does the East bar represent $7M, $8M, or $9M? This confusing quality of 3D bars is one of the topics in the next chapter.)*

on the data, the chart's business message will suffer. The chart might not be a lie, but despite the presenter's good intentions, it's certainly an awful shame!

These notions about geography also affect other elements of your presentation, such as photographs of plant locations, scenic backgrounds used for charts, diagrams of computer networks, and presentation of demographics (statistics by population group and by geographic region).

One Picture Says It All

The graphic in Figure 3.14 summarizes this chapter by committing at least three serious errors. First, the increase in sales is shown as horizontal bars, sacrificing the intuitive power of up-down. Second, the vehicle in the picture is traveling from right to left, a retrograde path that conveys the *opposite* of a feeling of progress (not a good technique when you're talking sales).

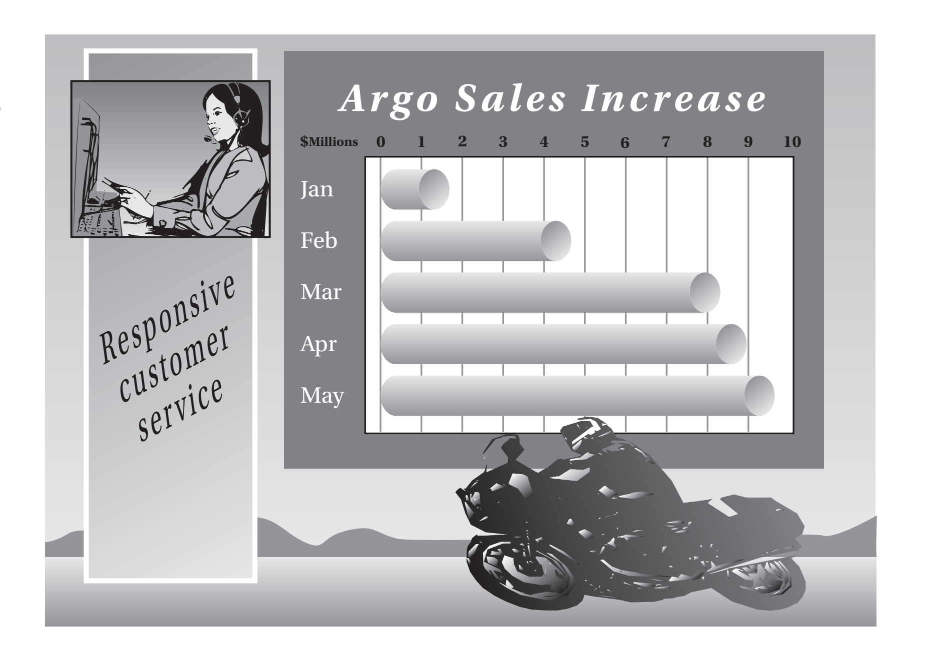


Figure 3.14 *Guess what's wrong with this picture and win a trip to the next chapter!*

The third error is one I haven't covered yet, and it's actually a compound error. Notice that a photo of a customer service representative has been inset near the graph. Now, to us intelligent primates, human faces are the most interesting thing on earth (despite what you might think about attractive bodies with little or no clothing). The face, in itself, is a major distraction from the intended message, simply because it is the wrong place for the audience to be looking. Furthermore, the face in the photo is looking to the *left* (not good), and (worst of all) her gaze is off the edge of the chart, away from the real subject of interest—the sales data!

Consciously or not, when the audience sees this chart, they will look at the face first, and then look where the face is looking. In the moment it takes to focus back on the bar graph, the presenter might have made his point and moved on to the next chart! (What was that he said?)

The next chapter gets truly serious about *xy* charts, which, as I've said, capitalize on our prejudices about orientation.

4

XY Charts

Or, The Old Story of Money and Time

PERHAPS the most common scheme for business charts is a plot of quantity versus time, as shown in Figure 4.1. Amounts—of dollars, say—increase or decrease along a vertical scale, or *y axis*. Time progresses from left to right along a horizontal scale, or *x axis*.

Many people assume they know how to read these charts because this up-down, left-right arrangement is so intuitive. In Chapter 3, I describe how this *xy* plotting scheme relates to meanings we have given to directions in space.

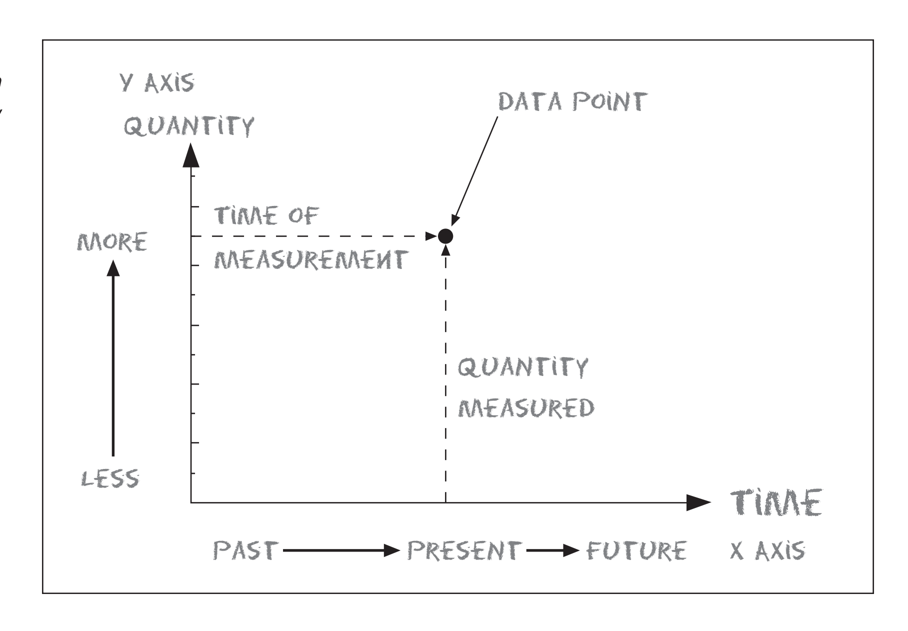


Figure 4.1 *People in Western countries seem to understand intuitively that up-down means* amount *and left-right means* time.

Thank You, René Descartes

The *xy* charting scheme is derived from the Cartesian coordinate system developed by the seventeenth-century French mathematician and philosopher René Descartes (whose name, by the way, can be translated "René of the Charts"). If you took trigonometry in school, you've probably had it with old René, but here's one of those dry subjects that can finally help you persuade people and make money!

In the Cartesian system, the position of a point in two-dimensional (2D) space is determined by two numbers, or *coordinates*, as shown in Figure 4.2. Each coordinate is a distance from a central point, or *origin.* A point in 2D, such as on a sheet of paper, can be located *x* numeric units horizontally and *y* units vertically from the origin. So drawing a line requires at least four numbers—two for each of its endpoints.

A 2D drawing or chart in the Cartesian system is flat, or *planar.* To add depth, a third dimension, or *z* axis, can be used, as shown in Figure 4.3. In 3D, a point in space is defined by three coordinates—*x, y,* and *z.*

In a different scheme called *2½D* representation, the *z* axis is used—not to add depth or solidity to a chart or drawing—but to hold separate drawing planes, like a stack of clear sheets (Figure 4.4), or *overlays*. In 2½D, the *z*-axis value is the *number of the sheet* on which an object is drawn. The order of the sheets determines the *visual priority* of objects. Objects on sheets with high visual priority cover up and hide the parts of other objects that lie directly beneath them.

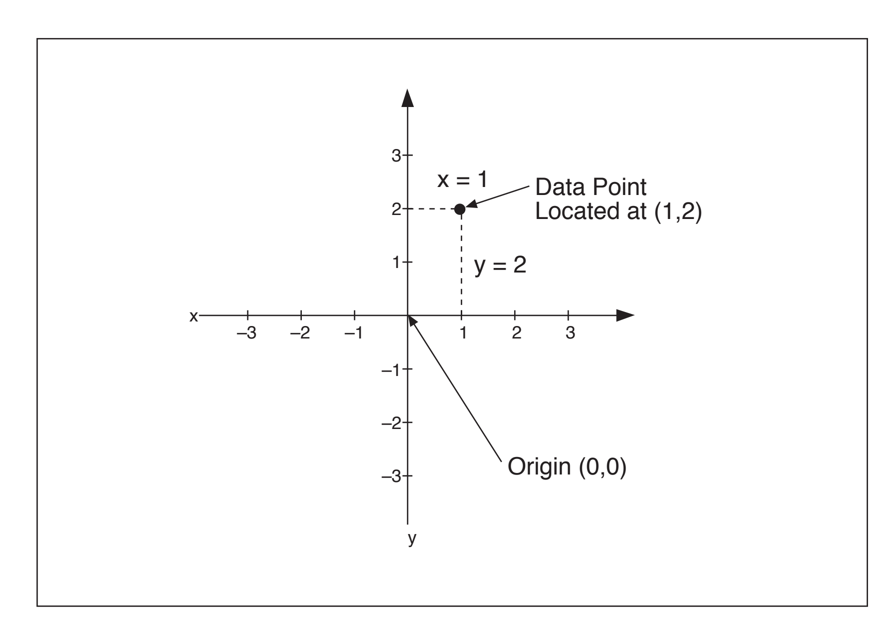


Figure 4.2 *The Cartesian system used for most charting is two-dimensional (2D).*

Figure 4.3 *Three-dimensional (3D) plotting of solids requires three coordinates for each point—x, y, and z.*

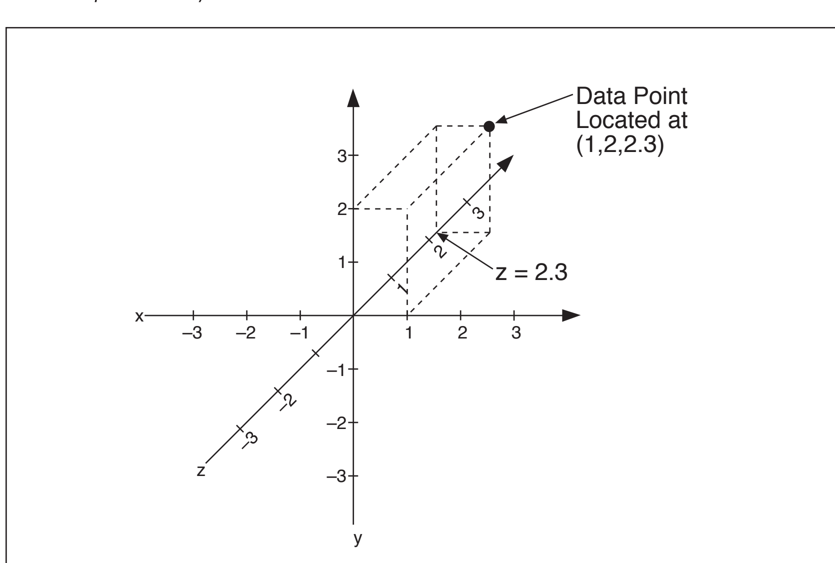
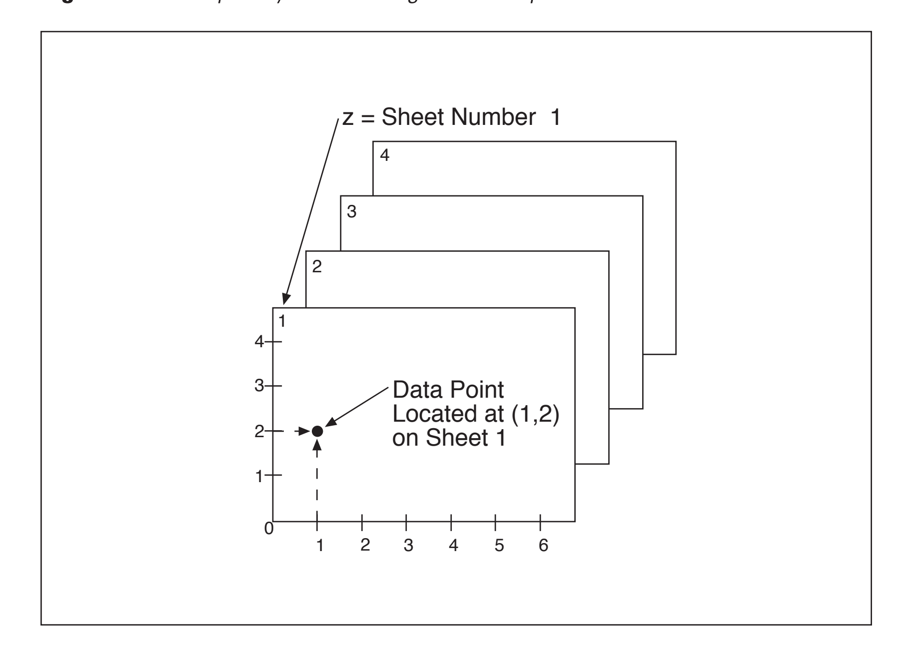


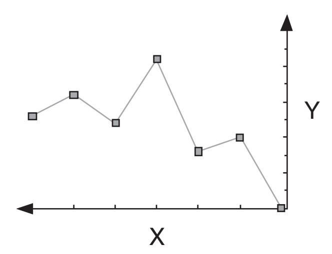
Figure 4.4 *Multiple-layered drawings can be represented in 2½D.*



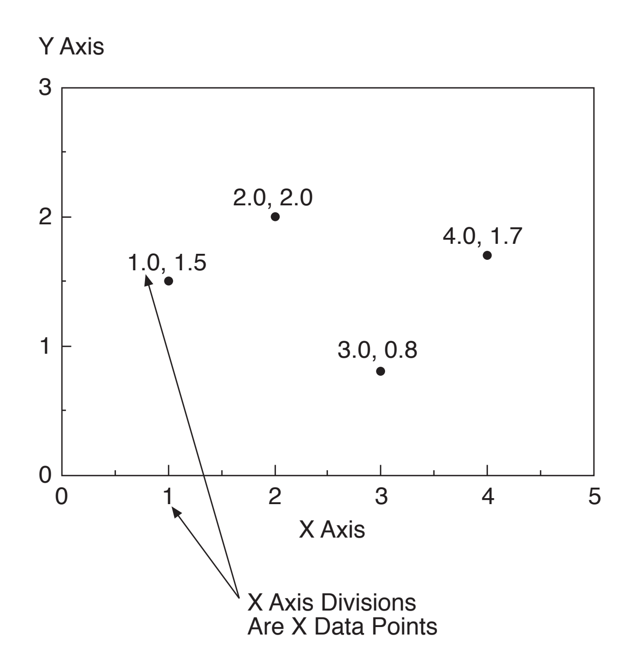
$$$$INVESTOR'S TIP$$$$

In trigonometry and scientific analysis, a Cartesian plot has four quadrants, extending to negative values of *x* and *y*. But in financial charting, there is only the single, top-right quadrant defined by positive values –*x* for time and *y* for dollars.

Especially for Western audiences, upward motion means gain or increase, and motion from left to right conveys the progress of time. If Hebraic scholars—who read from right to left—had invented *xy* charting, a typical plot might look like this instead:



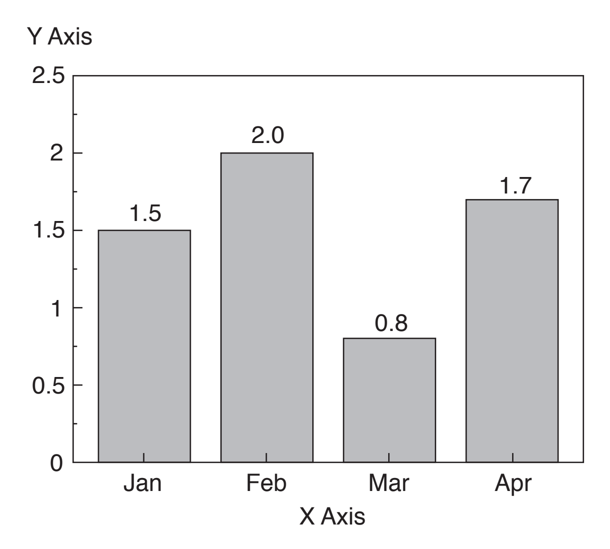
In the most common type of *xy* data chart used in business, a sequence of data values is plotted as a series of points on the chart in left-to-right order. The height of each point is its *y* value. Its *x* value is simply the next division of the *x*-axis scale:



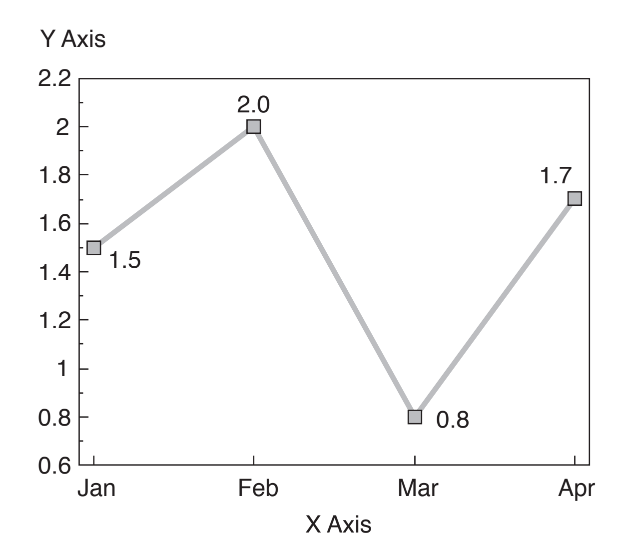


For most types of *xy* charts used in business, the numeric value of a data point is given as just the *y* value. The *x* values are the corresponding *divisions*—spaced at regular intervals—along the *x* axis. Often, x-axis divisions are marked by labels instead of numbers. For example, because the x axis usually represents the flow of time, the x-axis divisions might be months, labeled JAN, FEB, MAR, and so on. If you had three dollar amounts, such as 150, 185, and 210, these would be plotted in sequence as *y* values: 150 at JAN, 185 at FEB, and 210 at MAR. A chart type that does not necessarily locate points at regular x-axis divisions is the *scatter* chart, in which a pair of numbers (*x* and *y*) is required to locate each point.

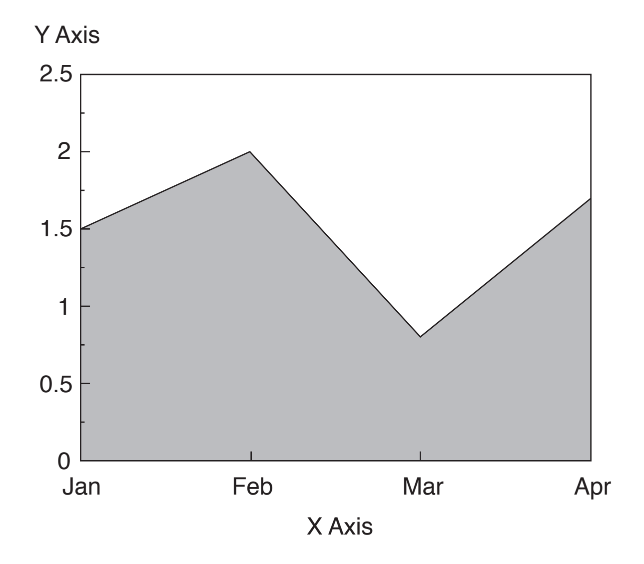
The sequence of points is called a *data series,* or *data set.* Lines or shapes are drawn to link the points in whatever plotting style you prefer. In these different plotting styles, data points are shown as heights of separate bars (*bar chart*),



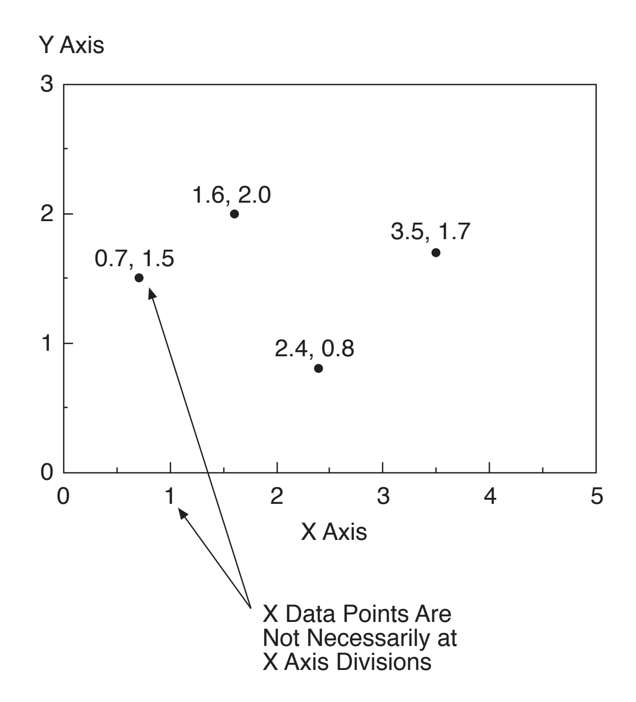
are connected to form a line (*line chart*),



or are connected with shaded areas underneath (*area chart*):



Data points can also be shown as dots or symbols in what's called a *scatter chart*, or *point chart*. In a scatter chart, the points aren't plotted sequentially. Two coordinates are used for each point, meaning that the *x* value can fall anywhere along the scale—not just at major divisions:





Scientists and statisticians—who labor meticulously to record each precious data point in lengthy experiments and surveys—just love those scatter charts. The idea is to throw the "raw" data up there and then look for patterns. The rest of us generally prefer something more digestible.

Some people use the term *xy chart* to refer only to scatter charts. In this book, I use it more generally to describe any of several chart types that have *x* and *y* axes, such as bar, line, and area charts.

These *xy* plotting styles can have many variations, including stacked and cumulative plots, and special formats.

Stacked or Cumulative?

$$$$INVESTOR'S TIP$$$$

For reasons explained in this chapter, area charts are likely to be extraordinarily confusing to financial audiences. Tell your story with stacked bars instead.

In an *xy* chart, sets of data can be built on one another in two different ways: stacked or cumulative.

Stacked Charts Are a Bumpy Ride!

In a stacked area or bar chart, each data set uses the previous set as its baseline: Each *y* value is added to the maximum *y* value of the previous set. So, where one set peaks, the next begins. In effect, each data set rides the bumps of those below it—built upon one another like bricks and mortar or layers in a cake. A set of stacked areas is shown in Figure 4.5.



It's not nice to stack the *lines* in a line chart. The notion of stacking won't be visually apparent, as it is with solid areas or bars. By convention, the baseline of each series in a line chart is usually assumed to be the line *y*=0 (the x axis).

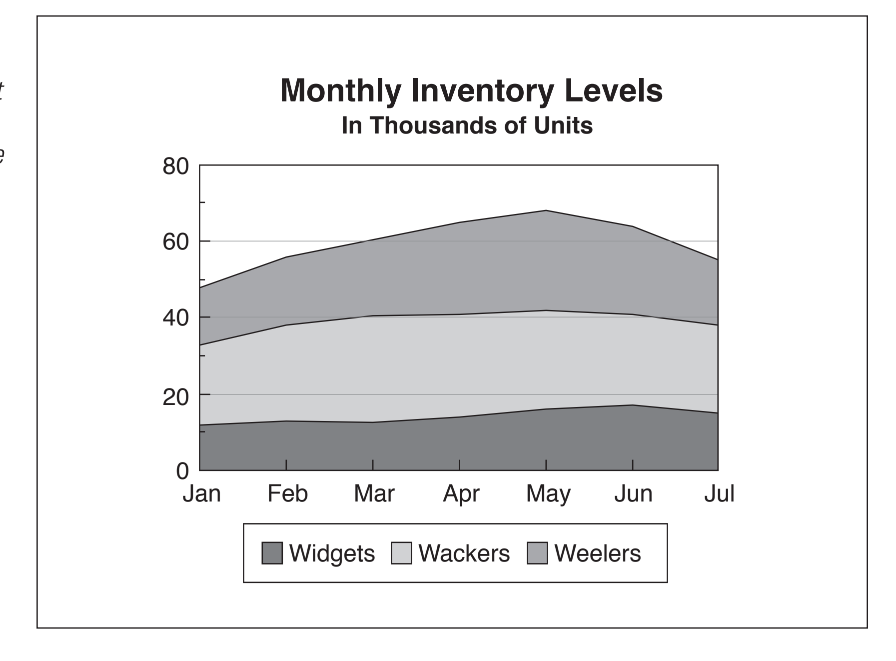


Figure 4.5 *Area charts are usually shown with* stacked *data series. That is, the starting point of one area is the top of the area below it. Putting the least "bumpy" data series at the bottom minimizes distortion of the other series.*

For bar charts only, an alternative to stacking several data series on the same chart is to *cluster* them. Separate bars are shown for each series at each division of the x axis (Figure 4.6).

Figure 4.6 *The clustered bars in the lower chart are an alternative to the stacked bars of the upper chart. Note the difference in the ranges of the* y*-axis scales—from 0 to 30 for clustered plotting, from 0 to 80 when the series are stacked on one another. Notice also in the stacked chart that the series with the least fluctuation has been placed on the bottom to minimize distortion of the other series.*



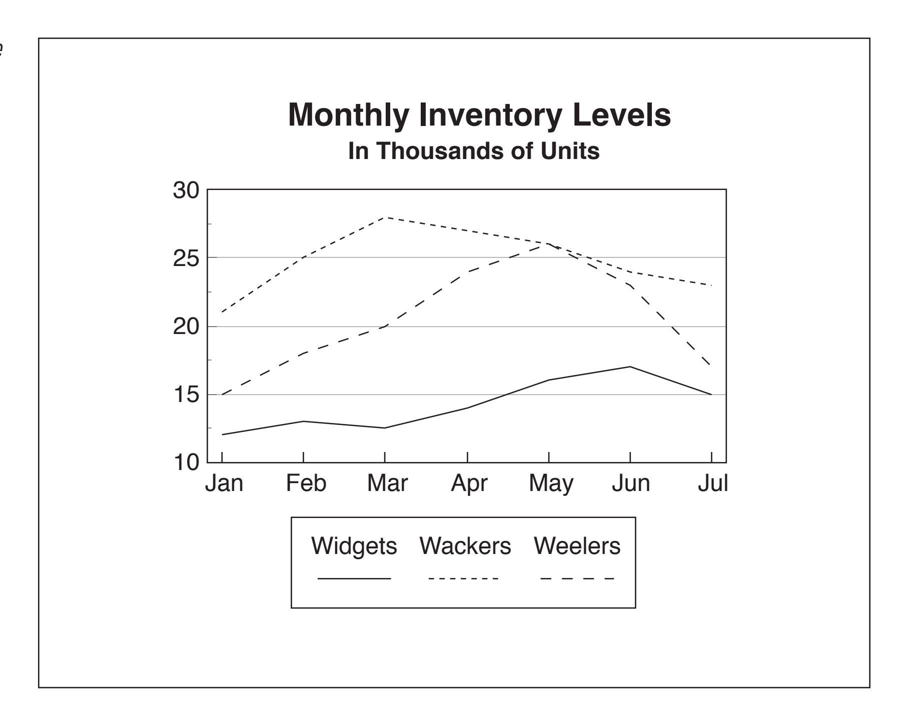


A clever liar can use this characteristic of stacked charts to his or her advantage: The fluctuations, or bumps, in the lower data series will enhance and exaggerate the bumps in the upper ones.

The bumpy effect will be most pronounced in area charts but can also affect perceptions about stacked bar charts. So, if you want to exaggerate the fluctuations in the data, put the bumpy series at the bottom (look back at Figure 4.6). Careful presenters who strive for honesty, however, will put the flatter data series on the bottom to minimize distortion in the other series.

As a general rule, area charts are plotted as stacked series, and line charts are not. Compare Figures 4.5 and 4.7, which both plot the same sets of data. In Figure 4.5, the areas are stacked, building on one another. The baseline of each series is the series beneath it. In Figure 4.7, which is a line chart, the series are not stacked, and each uses the *x* axis as its baseline.

Figure 4.7 *The baseline of all series in a line chart is assumed to be the x axis (y=0). (This chart plots the same data as Figure 4.5.)*



Cumulative Charts Can Be Confusing

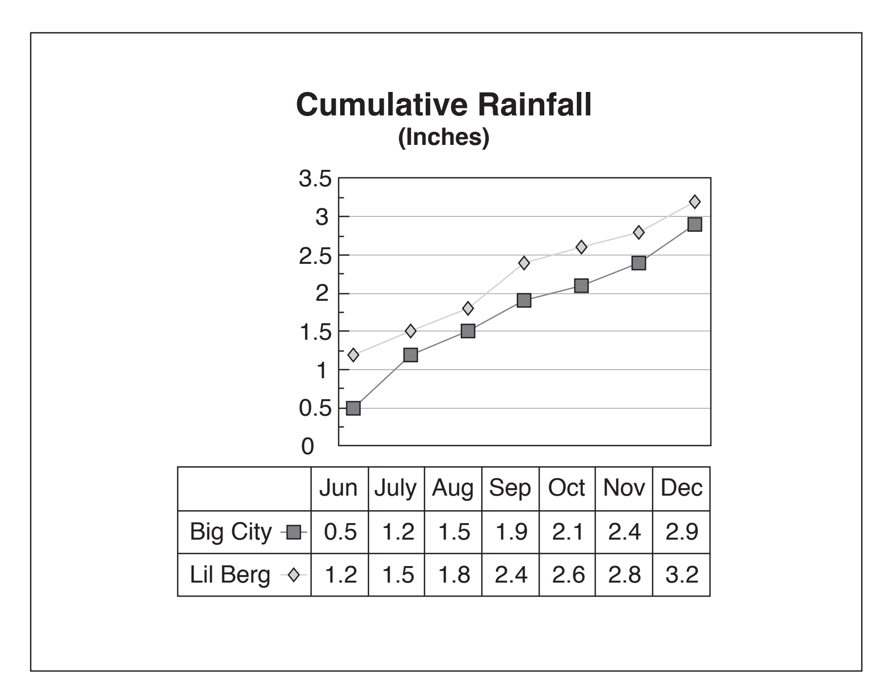
Figure 4.8 shows a different method of plotting. Each *y* value in a set—in this case, the rainfall in inches—is a running total of all the preceding values in the *same* set. This is the correct sense of the term *cumulative* when you're talking about charts.

Cumulative charts are a favorite with liars because the general impression is always optimistic: As long as the data points are positive numbers, the trend is always upward—even if some of the increases are small. In the rainfall example, September, October, and November were fairly dry months in both cities, but that fact is obscured by the upward slope of the cumulative totals.



Don't confuse *stacked* and *cumulative!* Stacking data series *does* make them appear to accumulate on top of one another, but this is not the meaning of *cumulative*. In a cumulative chart, the data accumulate only *within each series*. And, to avoid really confusing yourself and your audience, never mix stacked and cumulative plotting in the same chart! (Read more about it in Chapter 7.)

Figure 4.8 *In a cumulative chart, each point within a data series is a cumulative total of all preceding points in the same series. Cumulative charts look optimistic by nature.*



	Jun	July	Aug	Sep	Oct	Nov	Dec
Big City	0.5	1.2	1.5	1.9	2.1	2.4	2.9
Lil Berg	1.2	1.5	1.8	2.4	2.6	2.8	3.2

A careful presenter will only use cumulative charting for quantities that actually accumulate. For example, you could use a cumulative chart to show the growth of the principal in an investment with compounded interest. In such an investment, interest from the prior period is added to the principal amount before interest for the current period is calculated.

Special Charts for Special Occasions

Some special ways of plotting *xy* data are unique to specific fields or industries. These include *high-low-close-open (HLCO)* charts and *Gantt* charts.

Follow Stock Prices This Way

$$$$INVESTOR'S TIP$$$$

There's much more on HLCO charts in Chapter 10, Technical Analysis Basics.

Stock prices are typically shown as *floating bars* in a format called *high-low-close-open (HLCO)*. Each bar shows the high and low prices for a stock on a given day. Open and close—the prices at the opening and closing of the day's trading session—are shown as small horizontal *ticks* in the bars. Here are two different styles of HLCO floating bars. The thin one on the left is also called a *bar,* even though it's just a segment of a line.

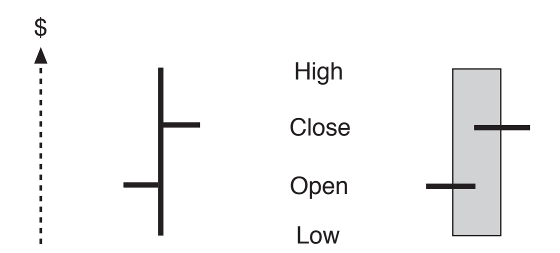


Figure 4.9 shows another HLCO chart. In this example, the number of shares of a stock traded each day is plotted as a separate set of bars at the bottom of the chart.

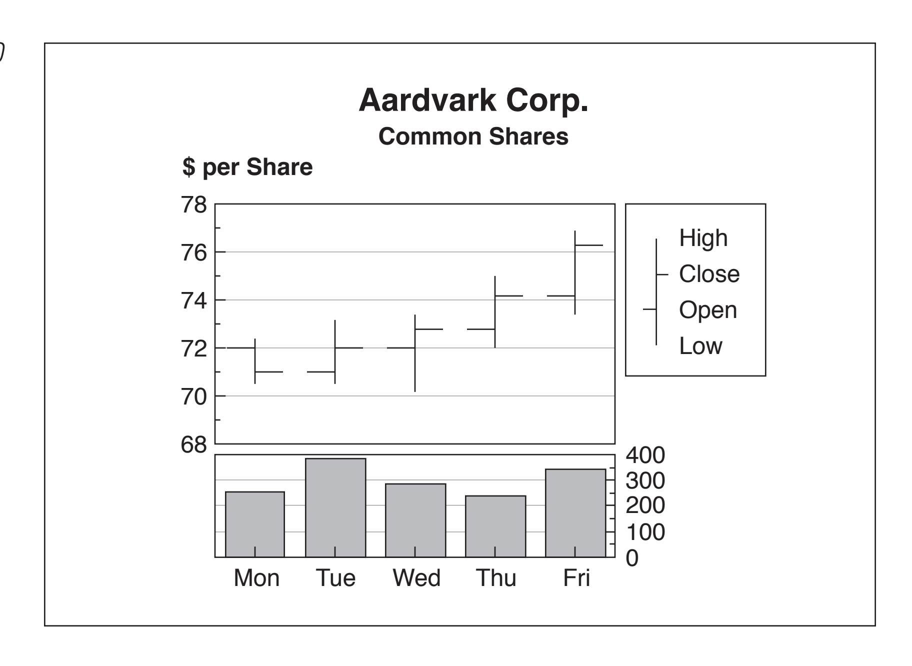


Figure 4.9 *In this HLCO chart, daily volume of this stock is plotted as another data series at the bottom.*



HLCO charts can be distorted by playing with the appearance of the bars (skinny or fat) and by manipulating the *x* and *y* axes. Warfare using axes is a big topic I'll defer to Chapter 6.

Project Managers Love to Stretch Mr. Gantt

As I point out in Chapter 3, another common use of floating bars is to represent time spans, or *durations*. In a *Gantt chart*, time flows in the familiar direction—from left to right—but the divisions of the *y* axis represent not quantities, but the different phases of a project or larger time span (Figure 4.10).

Project managers routinely use Gantt charts to show work schedules. Progress against a schedule can be highlighted by using two sets of floating bars: Planned and Actual. Any gap between the planned and actual bars represents a visible *slippage* of the work schedule.

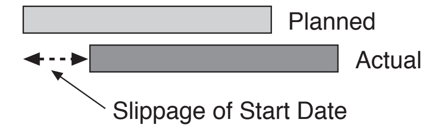


Figure 4.10 *In the Gantt charts used for project management, durations of work activities are plotted as horizontal bars.*

Software Development Project

ID	Task Name	Start	March				April					
			03/07	03/14	03/21	03/28	04/04	04/11	04/18	04/25	05/02	05/09
1	Design	Mar 15										
2	Review	Apr 5										
3	Test	Apr 19										
4	Debug	May 3										

Notice in Figure 4.10 how the interpretation of the floating bars is helped by the notion of time flowing from left to right. This impression would be defeated if the same durations were plotted instead as conventional, vertical bars, as shown in Figure 4.11. When the audience sees the vertical bars, the impression will be that the project is static, not making progress, even though the data in both cases is the same.

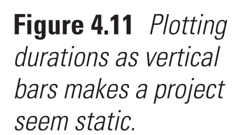


Project managers can lie with Gantt charts. If the *x* axis is elongated, the audience will get the impression that the overall time span has been lengthened. (All project managers want more time because they are always up against tight deadlines!)

The Lesson of the Grasshopper and the Gantt

Gantt charts are named for Henry Gantt, an American industrial engineer of the early twentieth century. Gantt described business organizations as mechanisms that are subject to three basic managerial functions: planning, organization, and control. Features of his charting technique have been designed to support each of those functions. Gantt's theories were studied by the young W. Edwards Deming, whose own work in quality control was taken much more seriously in postwar Japan than in his own country, the United States. Some historians believe that adopting Deming's methods was a major factor in Japan's stunning success in high-tech manufacturing.

The lesson? Study the Gantt, thou sluggard, and grow wise!



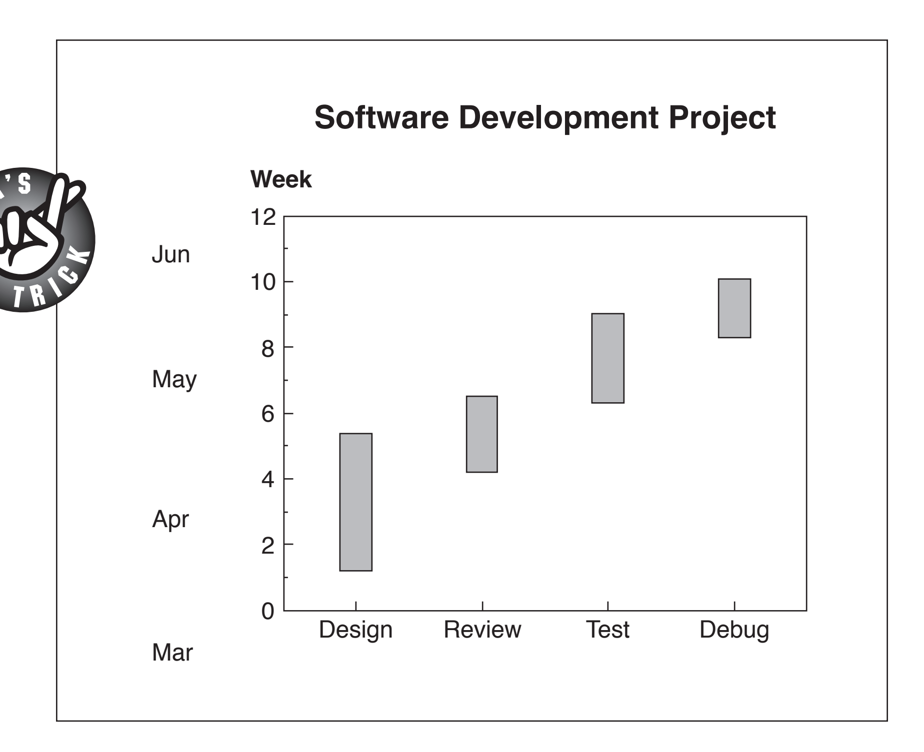


Figure 4.11 *Plotting durations as vertical bars makes a project seem static.*

Hanky Panky with X and Y

There are some cardinal sins you may want to commit, or not, when designing *xy* charts. Here are two of the liars' favorite tricks: First, turn vertical charts into horizontal format, subverting the intuitions of the audience about quantity (up-down) and time (left-right). Second, mix proportional and quantity-time relationships in the same chart. For example, use symbols instead of bars, which you'll see is a really nasty trick! (For an example, look ahead at Figure 4.13.)

Laying and Lying

Worried about the highs and lows in a vertical bar chart?



Laying a vertical chart on its side can be a form of subtle lying because you confuse the intuitive meanings of up-down and left-right. Figure 4.12 shows the same chart in both the conventional vertical bar

Figure 4.12 *Turning a conventional vertical bar chart on its side confuses the intuitive meanings of up-down and left-right.*

National Sales Volume by Region
(Millions of Units)

600
500
400
300
200
100
0
West North South East

National Sales Volume by Region
(Millions of Units)

West
North
South
East
0 100 200 300 400 500 600

LIAR'S TRICK

format and the horizontal format. Notice that the effect is quite different from Gantt charts, which also use horizontal bars but *preserve the flow of time from left to right.* In a horizontal chart, the positions of the axes have been transposed: The *x* axis—which usually represents time—runs up and down, and the *y* axis—which usually represents quantity—runs from left to right.

The horizontal format is misleading because the audience will need a moment—however brief—to adjust to the counterintuitive layout. An accomplished liar can use this slight delay in comprehension to achieve nefarious purposes. For example, the unscrupulous presenter might say something in that moment about the chart that he or she doesn't want the audience to fully understand: "These numbers are hypothetical, of course."

Horizontal bars used to represent quantities are misleading for another reason. They will be perceived as if they were *dependent on time*, and progressing toward a goal. The longer bars, which would be taller on a conventional vertical chart, will be seen as *closer to the goal*. In reality, there might not be a goal, or the goal might not be relevant to the chart.

Honey, I Blew Up the Data!

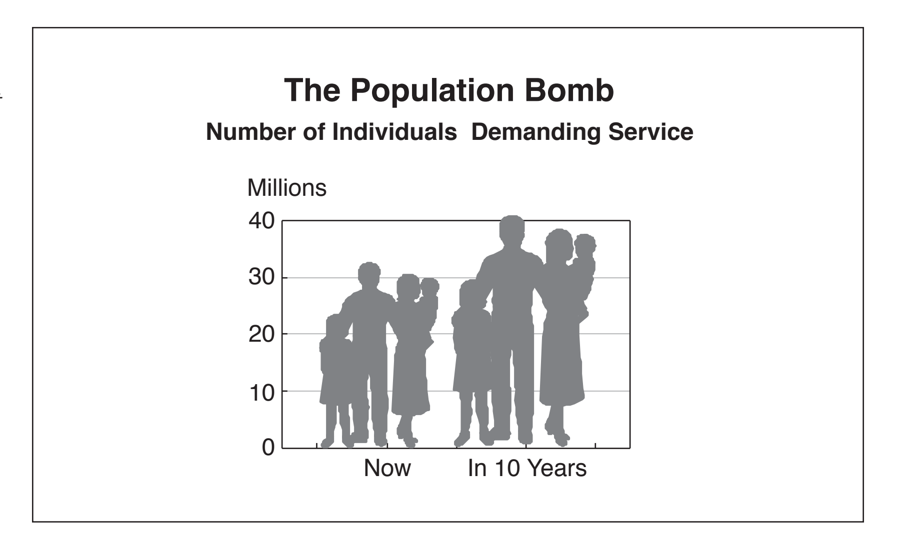
Recall in Chapter 2 that I describe pie charts as a means of showing proportional relationships, such as percentages or ratios. *XY* charts, on the other hand, are designed to represent actual quantities.

It's a recipe for disaster to mix proportional and actual relationships in the same chart! The designer of the chart in Figure 4.13 apparently thought it would add a creative touch to use symbols of people instead of bars.

The chart with the people symbols is a *huge* distortion of the data! This distortion arises from the difficulty of estimating the areas of objects visually, just as with proportionally-sized pies. (The gender stereotypes in this illustration might also preoccupy the audience, but that's a different issue.)

In the example, the projected increase from this year to next is just 20 percent. The designer of the chart has made the people on the right 20 percent taller, as though they were bars. However, the shapes on the

Figure 4.13 *The people on the right are 20 percent taller (the correct comparison) but are* 200 percent larger *by area!*

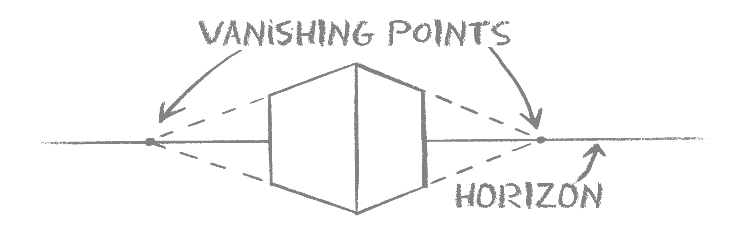


right—which have increased in width as well as height—are actually *twice* as large, by area. In other words, the increase is made to look much bigger than it really is: *200 percent instead of 20 percent!*

Tricks of Perspective

A favorite trick of liars who hang out in bars is to use 3D to make all the bars in the chart appear taller. In general, there's a right way and a wrong way to add depth to a bar chart. The difference lies in a trick of *perspective*.

In a 2D perspective drawing, the sides of objects appear tapered, converging toward one or more points on the horizon called *vanishing points*. A common type of perspective drawing uses two vanishing points, one at each end of the horizon:



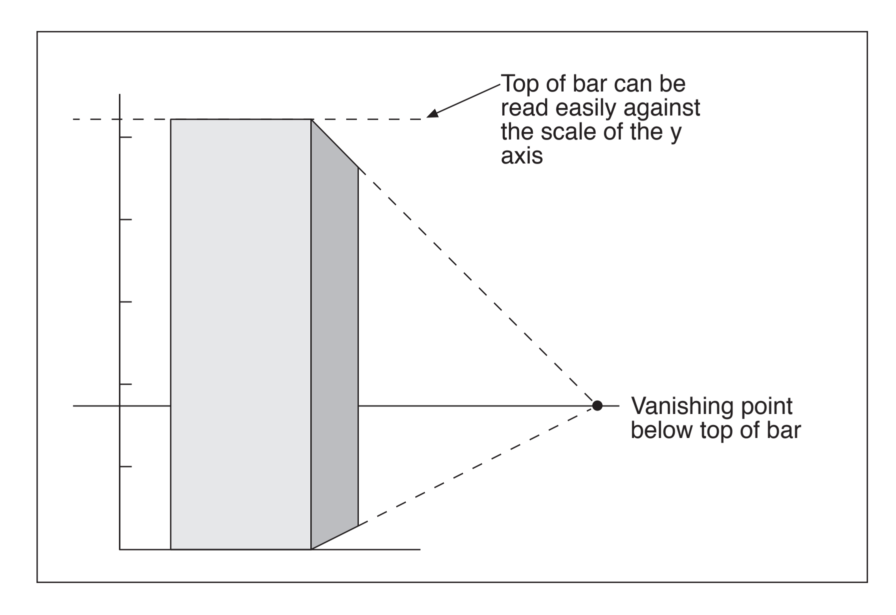


Figure 4.14 *Careful presenters put vanishing points* below *the tops of 3D bars.*

$$$$INVESTOR'S TIP$$$$

Let's be clear: Using any type of 3D chart in a financial presentation is probably a bad idea.

The rule for drawing perspective bars is simple: Careful presenters put the vanishing points *below* the tops of the bars, as shown in Figure 4.14. This makes the "front face" of each bar its tallest part, and it's obvious where the top of the bar is in relation to the scale.

The liar's trick with 3D is to put the vanishing points *above* the tops of the bars. This emphasizes the back edge of each bar, which appears higher than the front. This technique is a trick of perspective because the scale division also runs diagonally: The lower front part of the bar is just as tall as the back part. The trick makes it more difficult for viewers to estimate the heights of the bars accurately. The effect is shown in Figure 4.15.

If you plot multiple sets of 3D bars on the same chart, you might have no choice but to place the vanishing point above the tops of the bars, creating the potential for confusion shown in Figure 4.16. If you put the vanishing point lower, the bars in back might not be visible. If you place the vanishing point above, consider showing the numeric data as labels at the tops of the bars or in an accompanying table.

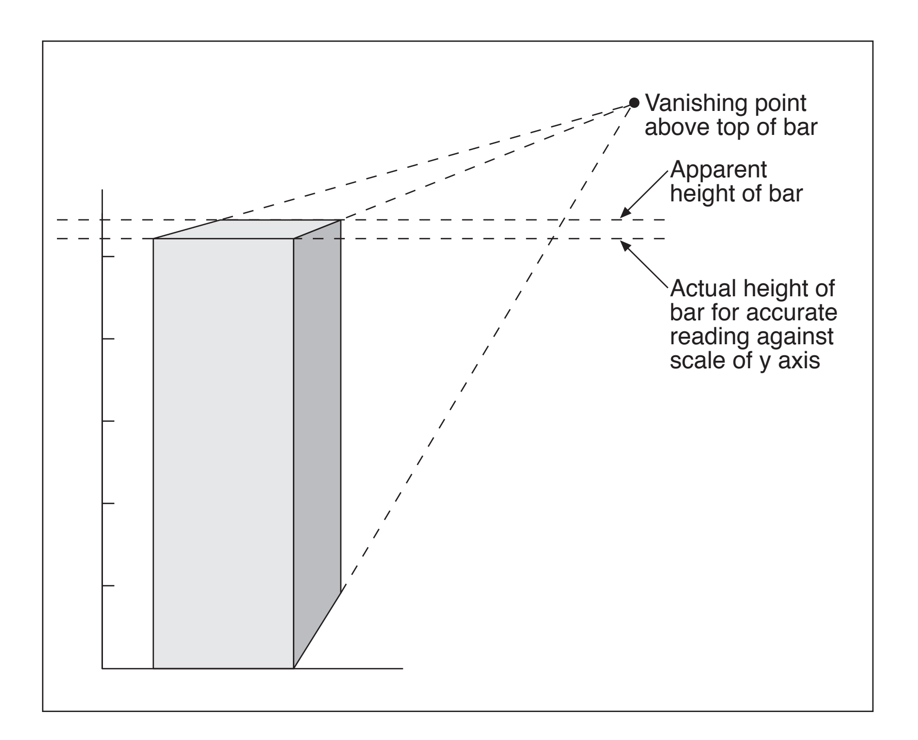


Figure 4.15 *It's a liar's trick to put vanishing points* above *the bars so that the back edges appear higher than the front.*

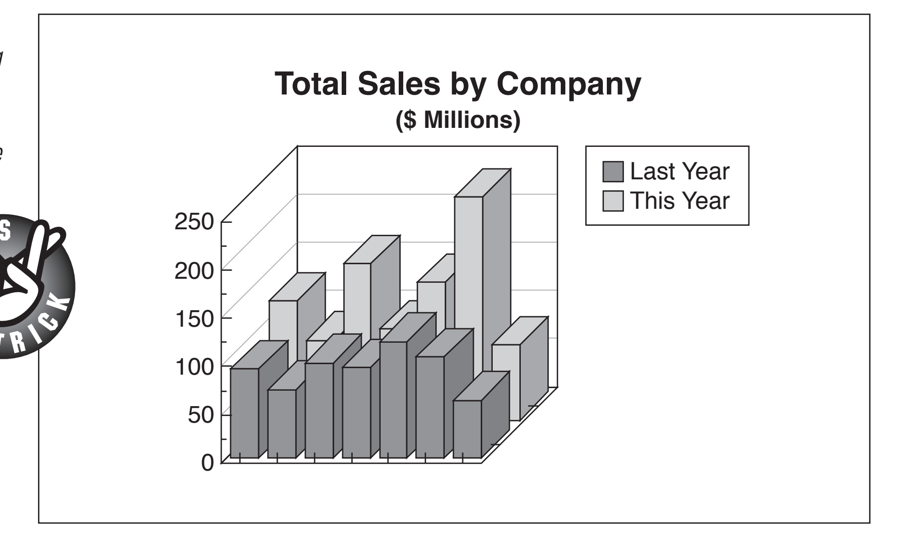


Figure 4.16 *In this set of 3D bars, the vanishing point is* above *the bars, making it difficult for the audience to estimate the true heights.*

An Exception to the Golden Rule of X and Y

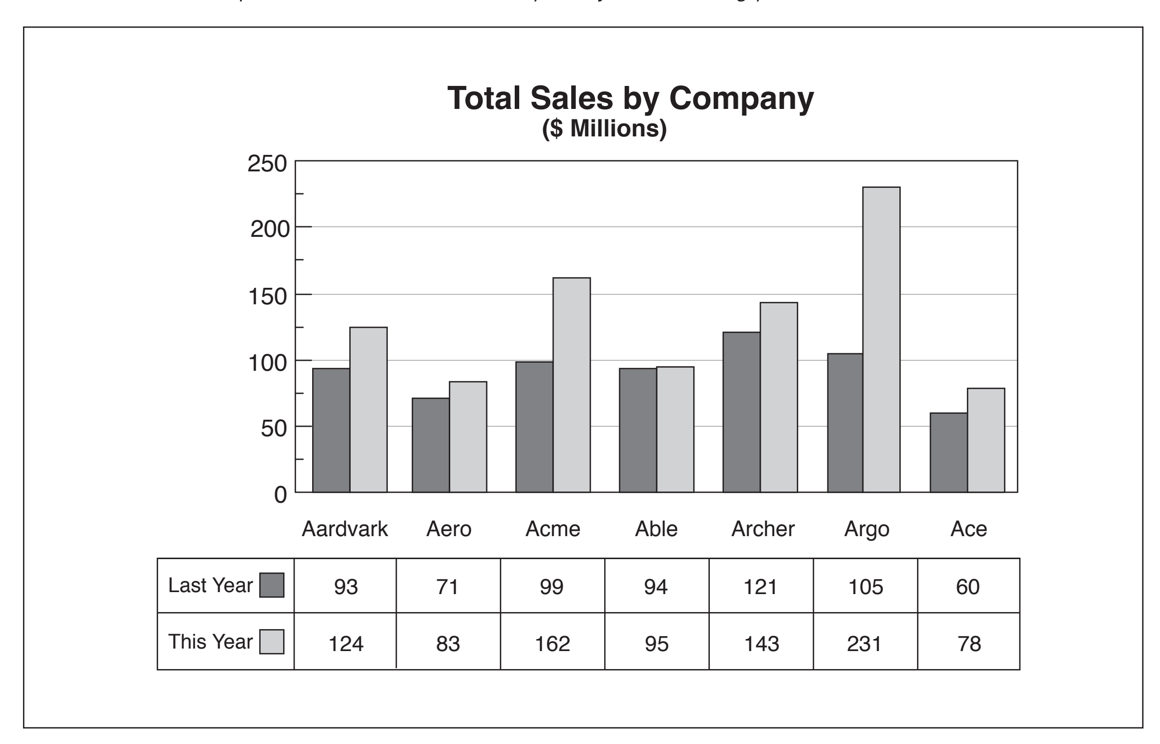
In an *xy* chart that strives for honesty, the vertical *y* axis should be an amount and the horizontal *x* axis should indicate the flow of time from left to right. To do it otherwise invites misunderstanding, *with one exception.*

In some bar charts, such as the one in Figure 4.17, divisions of the *x* axis are not time spans, but labels of different categories.



Even if the *x* axis has nothing to do with time, the audience won't be able to avoid associating left-to-right movement with *progress*. If you want to exploit this tendency, put the set of bars you want to emphasize on the *right end* of the chart. Liars may want to shift older, perhaps more favorable, data to the right end of the *x* axis.

Figure 4.17 *Even if the x axis does not represent the flow of time, viewers will regard the bars on the right as the most recent or important. That's not necessarily bad, just something you should be aware of.*



	Aardvark	Aero	Acme	Able	Archer	Argo	Ace
Last Year	93	71	99	94	121	105	60
This Year	124	83	162	95	143	231	78

Looking Ahead to Your Future as a Zen Master

The next chapter covers an Eastern cousin of the *xy* chart: the *radar chart,* which looks like a mandala, if you want to get spiritual about it. You may be surprised to learn that for the same reasons that Western audiences prefer the familiar *xy* chart, the multifaceted radar chart is a great persuader in Japan!

5

Radar Charts

Or, Watch Out for That Zen Speed Trap!

A **GRAPH that looks like a geisha's parasol** without the silk is called by the less picturesque name *radar chart*—a term borrowed from the target-shaped screen used by air-traffic controllers.



I can't resist the image of the Japanese parasol, because I've heard that these radar charts are particularly popular in Japan. The reasons have much to do with the differences between the expectations of Eastern and Western audiences—but let's discuss the basics first.

Axes, Grids, and Criteria

Figure 5.1 shows a typical radar chart—a rating of a new car model according to five criteria: fuel economy, handling, acceleration, styling, and ride. Like the spines of an umbrella or the spokes of a wheel, several axes radiate outward from a common center. The optional gridlines are shown as concentric circles. Each gridline marks a uniform distance from the center. In this case, the distance from the center on a gridline is the car's score for that criterion.

Each axis can represent a separate type or category of measurement. (The mathematicians call it *multivariate analysis*.)



If you want to be able to draw conclusions from the *shapes* of the data plots, the scoring must be consistent for all criteria. For example, let's say you rated a company's customer service on promptness, courtesy, accuracy, effectiveness, and follow-up. So, if you score courtesy on a scale from 1 to 10 using precision to tenths of a point (6.8), you should score the other criteria exactly the same way. Liars don't necessarily follow this rule, for reasons that should be clear at the conclusion of this chapter.

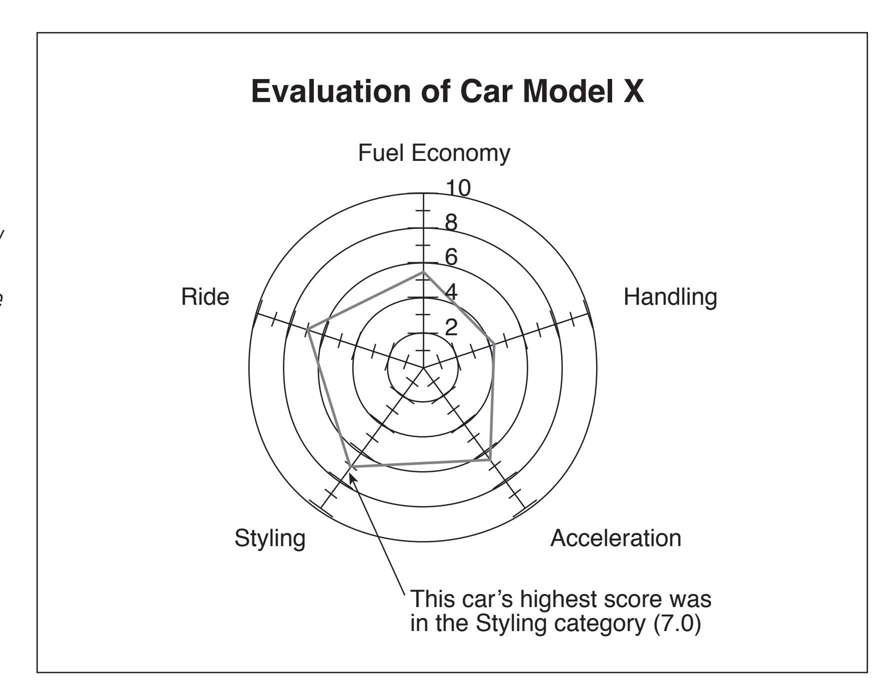


Figure 5.1 *The axes and gridlines of a radar chart let you describe a thing graphically according to multiple criteria—all in the same chart. In this case, a new car model is evaluated on a scale of 0–10 in five criteria: fuel economy, handling, acceleration, styling, and ride.*

A Delicious Example

This multivariate stuff is best explained by example. Continuing in the Japanese mode, let's assume that I'm the emperor's food taster. It's my job to sample the emperor's meal before he takes a bite—to make sure he isn't poisoned, yes, but also to advise him whether a dish is appetizing and worthy of his discriminating palate.

In my long and lucky years as the emperor's most trusted surviving taster, I have developed the habit of rating each new dish based on five key criteria: appearance, aroma, taste, texture, and digestibility.

In each of these categories, I give each dish a numeric rating from 0 (disgusting) to 10 (most worthy). Here is a recent memorable meal, which I have set out as an attractive table.

	APPEARANCE	AROMA	TASTE	TEXTURE	DIGESTIBILITY
MISO SOUP	6.2	5.9	6.0	6.1	6.2
CUCUMBER SALAD	4.5	2.0	4.2	6.5	6.0
TUNA SURPRISE	7.0	6.5	5.3	3.2	4.1
COOKIES	3.4	9.5	9.8	5.6	2.6

$$$$INVESTOR'S TIP$$$$

Radar charts might not be appropriate for showing financial results. But they're just the thing for focus groups and market surveys.

Each of the dishes, in this case, is a data set. Figure 5.2 shows how one of the data sets—Miso Soup—can be plotted on the same radar chart grid you've already seen. The numeric values in the data sets are plotted as data points *on* each axis (rather than between the axes, as in an *xy* chart). For example, the Soup course is marked at 6.0 on the axis labeled Taste. Points in the same data set are then connected to form a line plot. So, line segments connect the Soup scores among the five axes. For Soup, the result is a nearly regular pentagon because the rating for each category is nearly the same.

The outermost plots have the highest scores. You can see in the table that Cookies got the highest scores in Aroma and Taste. Notice, however, that the Cookies plot in Figure 5.3 is irregular in shape. Its pentagon looks squashed: Low scores in Appearance and Digestibility distort the plot.

Figure 5.2 *The plot of the data set Miso Soup is a nearly regular pentagon because it scored equally well in all criteria.*

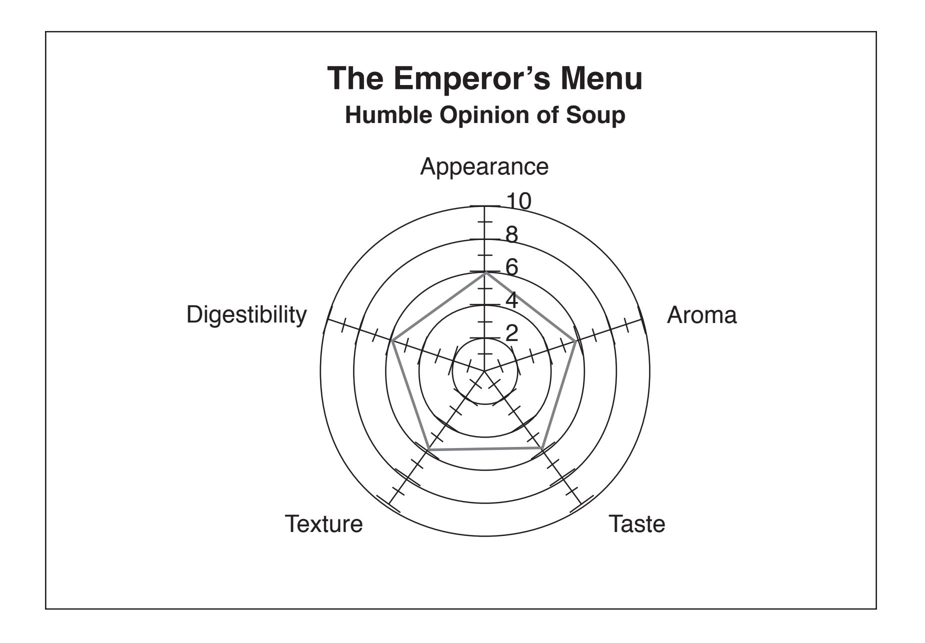
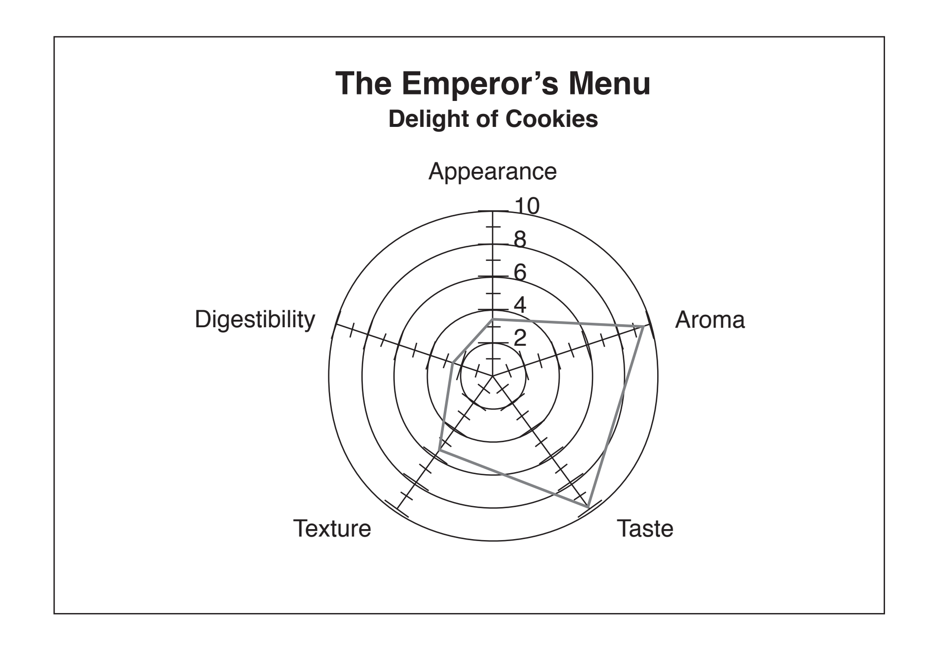


Figure 5.3 *The plot of the data set Cookies is irregular because it scored high in only two criteria—Aroma and Taste.*



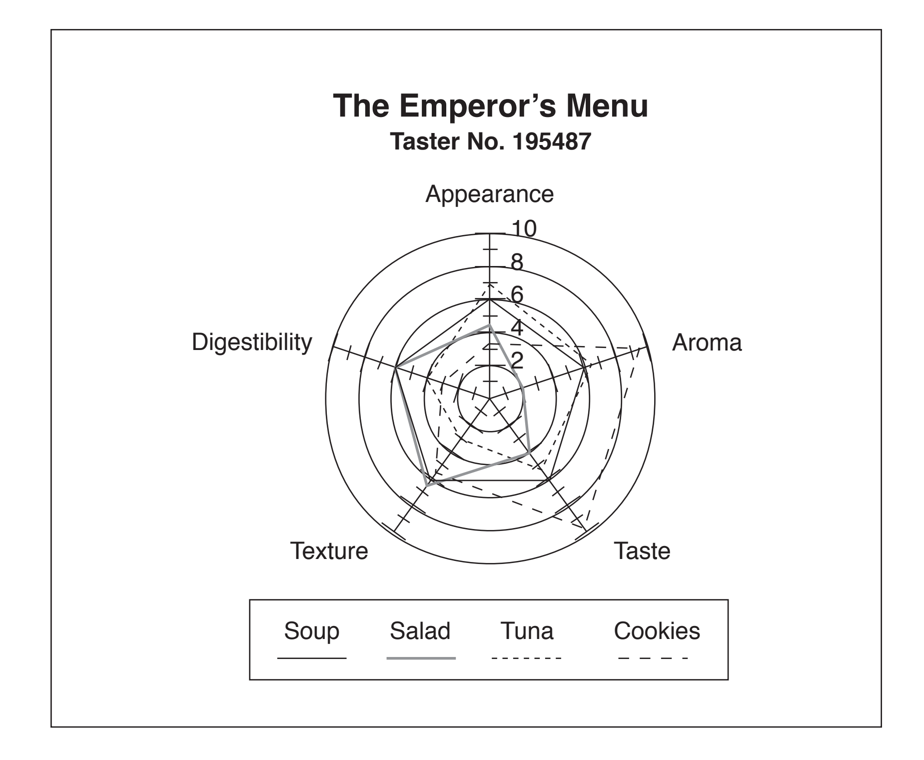


Figure 5.4 *The shapes—and merits—of the dishes can be compared when you plot them all on the same chart.*

The soup scored much lower than the cookies in Taste and Aroma, but it had uniformly good scores in the other categories. As a result, its plot is regular, or symmetrical. Soup is therefore not the most spectacular, but it is certainly the most "well-rounded" dish, both in terms of overall scores and its symmetrical plot on the radar chart.

The scores of the various dishes can be compared at a glance by plotting them all on the same set of radar axes, as shown in Figure 5.4.

A Zen Quest for Symmetry

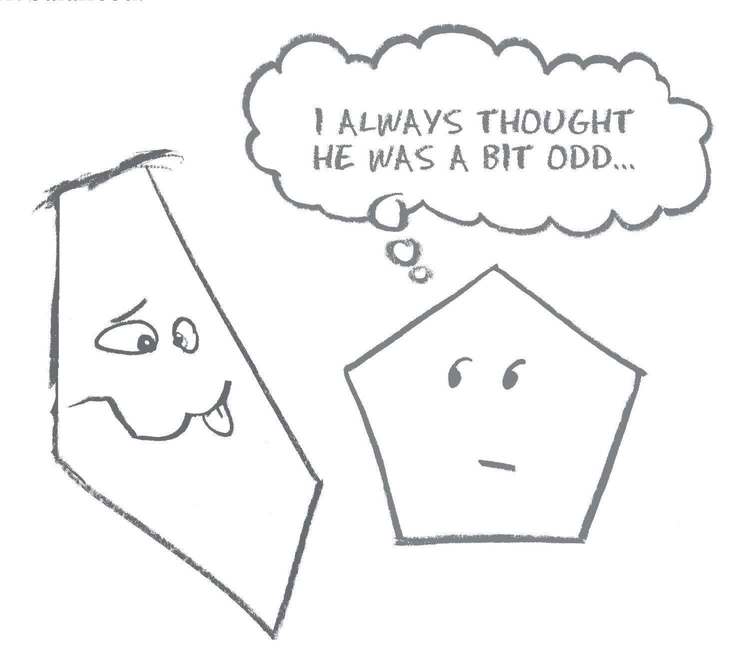
The quest for symmetry is the beauty of the radar chart. Its format implies a desire for balance and harmonious composition. The data set with the most nearly regular shape will be the most well-balanced among the various criteria. Regular shapes therefore represent

optimized choices, equalized trade-offs, best overall values. By contrast, the more distorted the shape of a radar plot, the more it will be skewed in favor of one or more of the criteria, and so will have a less desirable mix of scores.

So, the radar-chart format imposes a bias on the interpretation of the data: Symmetrical plots are "good," asymmetrical "bad." This bias is just as strong—although decidedly different from—the up-down, left-right biases inherent in *xy* charts.

Herein perhaps lies the reason for this chart's popularity with Japanese business managers. The conventional *xy* chart is rooted in the Western notions of progress (*x* axis goes left to right) and gain (*y* axis goes up and down). The radar chart, on the other hand, is literally centered. The impression is static, unhurried. The objective is balance, consensus, harmony. In this scheme of plotting things, the spectacular but erratic performer will always look a bit odd. The most pleasing shape will be presented by the competitor who is perhaps only adequate, but is nevertheless consistent in all areas.

Because the managers seek harmony and balance, the radar chart suits their mindset. They will be prone to reject alternatives that don't look balanced.

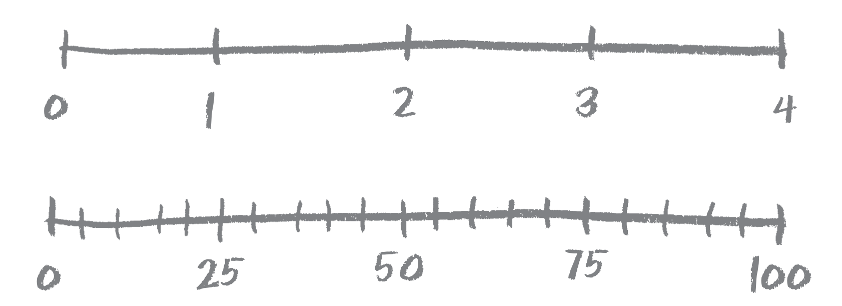


The Hazards of Navigation by Radar

In an honest radar chart, all axes should have the same scale. If axes have different scales, the symmetry of the plots, if it appears at all, will be meaningless.

Cheaters might pretend that the dissimilar scales on their radar charts are proportional to each other. For example, three spokes of a mixed-up chart might represent different scoring methods: 0 – 4, 0–100, and 1– 10. If the scales are truly equivalent, you should convert all scores so they can be plotted against comparable scales. If, however, there is really no relationship between the different scales, you shouldn't attempt to use them on the same chart.

Let's say you have two sets of criteria: One scale goes from 0 to 4, the other from 0 to 100. It might seem perfectly honest to stretch the first set of scores to match the second so that both sets can be plotted against the same scale. That is, you'd take a score of 4, which is at the top of the 0 – 4 scale, and *scale* it upward, or expand it proportionally-multiplying by 25. The result is 100, the same as the top of the 0– 100 scale.



However, the scores on the 0 – 4 scale are probably less precise than those on the 0–100 scale. You can see this if you try the conversion in reverse: Convert the 0–100 set of scores to the 0–4 scale by dividing by 25. Let's say you're scoring to the nearest tenth, or one decimal place, and you have a score of 85.6 on a scale of 0–100. Divided by 25, that

85.6 score becomes 3.424 on the 0–4 scale. Now, it's unlikely that you would have taken the pains to refine your scores to *three* decimal places on the 0–4 scale, so the two sets of scores still aren't truly comparable.

As a rule, you should convert all sets of scores in a radar chart to the *least precise* scale. In this case, you would divide the 0 –100 scores by 25 and plot them on the 0–4 scale rather than scaling the 0–4 scores upward. *You must then round all sets of data to the same number of decimal places.* So, if you convert 85.6 to 3.424, you must round the score to 3.4 so that it is comparable with the other 0–4 scores. You will lose the fine gradations of the 0–100 scores, but you never had them at all in the 0–4 set!

To be fairer yet, you should *truncate* the values, simply lopping off the extra digits. For example, a result of 3.467 rounded is 3.5, but truncated is 3.4. Using the truncated values makes a more valid comparison because the finer precision can't even be estimated in the coarser numbers, where it was never measured in the first place.

Here's Another Radar Trap!

A particularly nasty trick with radar charts is to intermix valid charts with questionable ones as adjacent pages or slides in the same presentation—that is, a liar might show some radar charts that have consistent scales alongside some that don't. The impression will be that all the charts are internally consistent.



The most dangerous liar's tricks with radar charts come down to scaling the data to fit some preconceived comparison—to make a plot seem either more regular or more distorted in relation to other data sets on the same chart.

But the tricks don't end there. Liars will try not only to scale the data, but they might also *scale the scales!* A clever chart maker can adjust and otherwise monkey around with the axes of a chart, the very device we honest types use to interpret the values of data points!

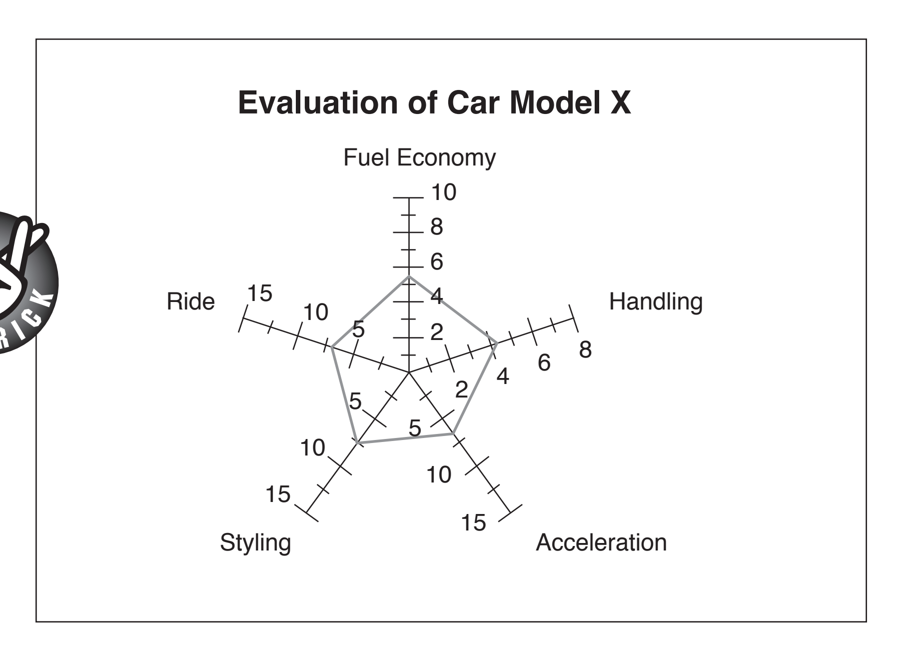


Figure 5.5 *Here's the same data I showed you in Figure 5.1, but with the scales adjusted to improve the shape of the plot.*

For example, I can manipulate the scales of a radar chart to make one of the plots seem more symmetrical, and hence, more desirable.

Look what I did in Figure 5.5. I took the data from Figure 5.1 and replotted it — this time, scaling each axis separately. Instead of requiring the scales to be consistent — each ranging from 0–10 — I've set some of them to different maximum bounds. In my new chart, the maximum extent of any scale is always *two divisions greater* than its data value. (There's no special reason for picking *two*—the choice just happens to work well with this chart.)

Remember — neither the data nor the scoring method has changed at all: The scores still range between 0 and 10. However, some of the criteria are now scaled 0–15, another 0–8.

Maybe because I can sound sincere, my audience might think I have some good reason for doing this, such as *weighting* the scores. Weighting multiplies scores in incompatible units by different factors so that they can be compared on the same scale. But the awful truth is — I have no legitimate reason at all for adjusting the scales!

My dishonorable reason for changing the chart was just to make the plot prettier. It is now more regular in shape, and, according to the hidden biases of radar charts, this car now looks like a better buy!

To shield you against such heinous deception, I cover manipulation of axes — a subtle vice that can corrupt *xy* and radar charts alike— in the next chapter.

6

Axes

Or, Handy Tools for Dismembering Your Enemies

$$$$INVESTOR'S TIP$$$$

Manipulation of chart axes is the most egregious fault in financial charting. And the default behaviors of PowerPoint and Excel don't make it any easier even when your intentions are good!

AS wielded by chartmasters, the singular of *axes* (pronounced *ak*-seez) is *axis*—not *ax*. But in the wrong hands, these axes can be just as dangerous. On an *xy* chart, an axis is one of two numeric scales on which the chart maker plots data points. Axes also help readers of a chart estimate the numeric values of its plotted lines, bars, or areas.

Be on your guard! Manipulated by an ambitious liar, chart axes are no crude weapon. Their potential abuses comprise a whole armory of hazardous devices. Learn about them now, or you risk being surprised under battlefield conditions!

To understand these techniques, you must learn to think like the enemy. Whether your actual battle statistics are grim or hopeful, you can control what your audience thinks about a chart just by changing its axes. You can adjust levels of panic or enthusiasm about your message almost as easily as those Roman emperors who sent gladiators to their fates with a brusque thumbs-up or thumbs-down.

Now, you can't always turn a negative into a positive—that favorite trick of every salesperson who gets a customer objection. But you can either exaggerate or minimize a chart's net effect. You can smooth a wildly fluctuating cost curve until it seems barely a ripple. Or you can crank up a modest increase in earnings so that it skyrockets—literally off the scale.

I strongly advise you to become proficient with the chart warrior's gallery of weapons. You might not feel obliged to use them, but being well prepared might discourage your opponents from using them on you!



Some of these tricks apply to radar charts, as well. *XY* charts are described in Chapter 4, radar charts in Chapter 5. This chapter discusses *xy* charts, because the examples will be clearer in this commonly used format. But remember that these types of manipulation can be applied to any chart type that has numeric scales. For an example of radar-chart scales that have been seriously abused, see Figure 5.5 in the previous chapter.

Pick Up Those Axes and Follow Me!

Now that I've cautioned you about the harm you can do, let's take a closer look at the use and abuse of chart axes. Strictly speaking, an axis is a line that indicates a direction in space, or *dimension:*



Presumably as a help for chart readers, the line can be marked off at regular intervals, or *divisions:*

Marks that indicate divisions are called *ticks*. As a further aid to measurement, big ticks can mark *major divisions,* little ticks *minor divisions,* much as you'd find on a ruler:

Of course, the ticks don't mean much unless you label their values. An axis with labeled divisions is called a *scale:*



The minimum and maximum values on the scale define its *bounds*. All the values in between make up the axis *range*.



Those clever mathematicians have a special notation—a type of *subscripted* variable—for axis bounds. Referring to the *y*-axis, the minimum value is written y_{min} and the maximum y_{max}. The values in a data series, when shown as variables, can be numbered sequentially by the same method: y_1, y_2, y_3, and so on. If you don't know how many values are in a series, the last value is y_n or the *n*th.

Okay, let's see what kind of trouble you can get into by adjusting the scales on a chart.

Bounding the Range Is a Good Exercise

One way to draw an axis is to begin just a bit lower than the least value in the data series that you intend to plot, and to end a bit higher than the greatest value. For example, if the *y* values in the data series ranged from 3.5 to 8, it might seem reasonable to start the *y*-axis scale at 3.0, and end it at 9.0, as shown in Figure 6.1.

Notice that the trend in this plot peaks quickly, then moves decidedly downward. Because it measures performance—the average number of orders taken per month by each salesperson—the audience might

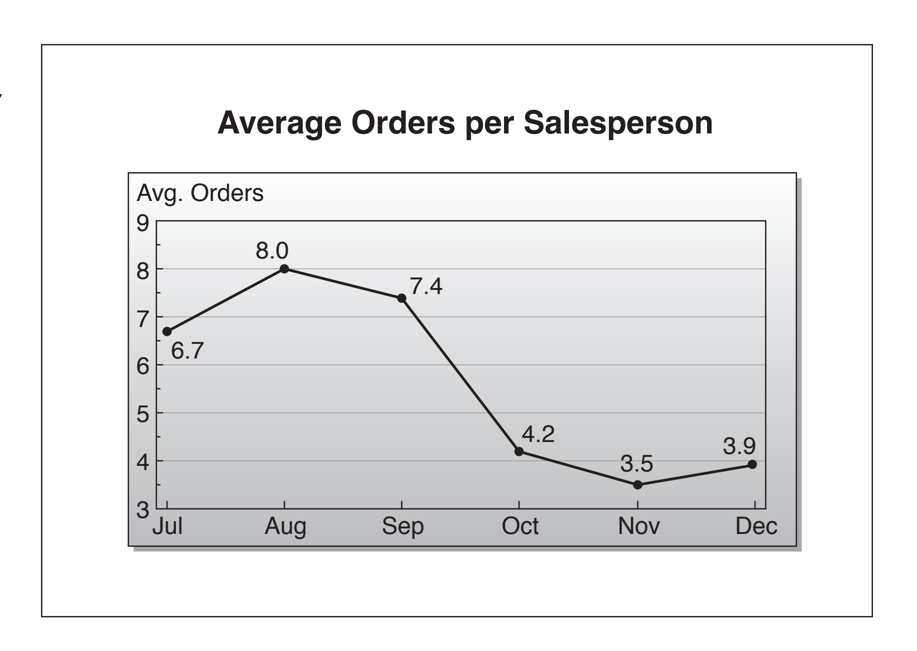
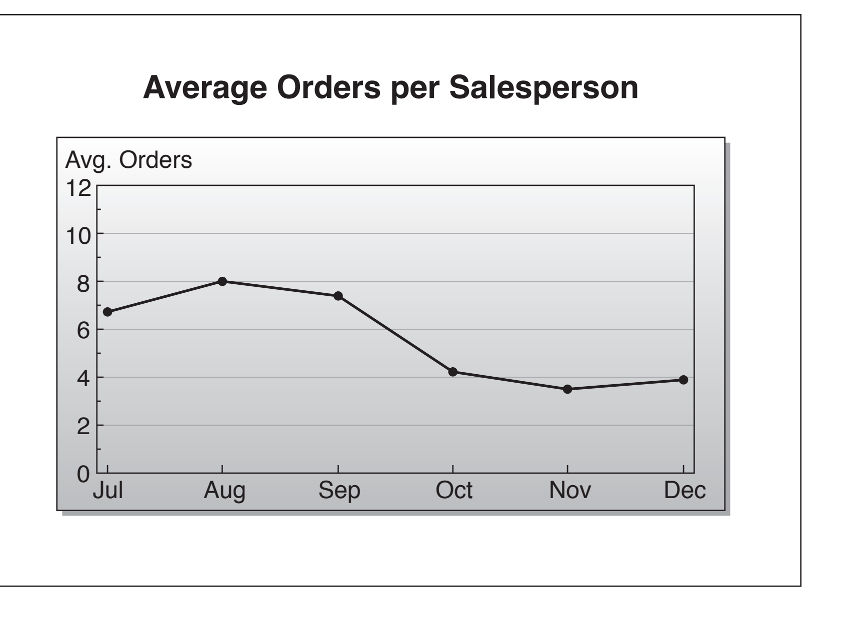


Figure 6.1 *The range of y-axis values is 3.5–8, and the bounds of this y-axis scale are 3.0 and 9.0.*

Figure 6.2 *Expanding the y-axis range here minimizes fluctuations in the data plot.*



grow gloomy watching the line plummet. But, let's say I'm one of those unscrupulous presenters. I want to soften the harshness of my bad news—and I'm not reluctant to cheat a little.

On my revised chart, I'll start the *y*-axis scale at zero—that's logical, isn't it?—and end it at 12. Here's the effect, as shown in Figure 6.2. While I'm at it, I'll omit the data labels.



To flatten a plot—minimizing its fluctuations—*increase* the *y*-axis range.

Whew! I guess my numbers aren't so bad, after all. As I show them this, I'll be saying that our sales are experiencing a "modest seasonal adjustment"—and I'll move quickly to the next chart!

The next chart shows the average number of sales orders taken per month by my staff. Although we're not the top performers in the company, I'm thankful to note that there's been a modest increase—from 4.5 to 4.7. I had good luck starting the *y*-axis range at zero before, so I'll make the scale from zero to 5, as shown in Figure 6.3.

Unfortunately for my future with the company, this plot looks awfully flat. Maybe starting at zero wasn't such a good idea. I'll make the axis

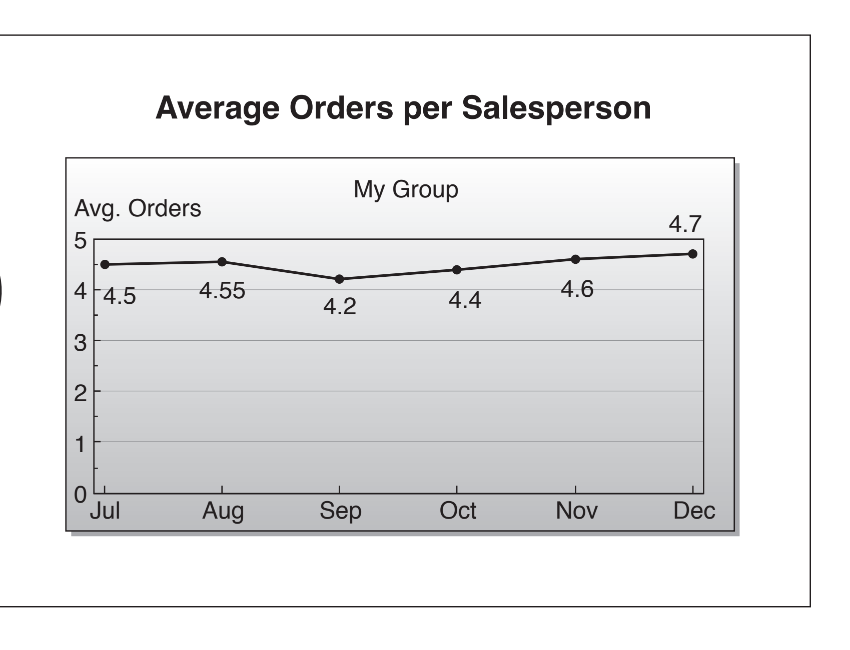


Figure 6.3 *Here's another example where expanding the axis range flattened the plot. Unfortunately, that wasn't the effect I wanted.*

range narrower—*really* narrow this time—starting at 4.15. And I'll end it at 4.6. My last data point will go right off the scale, as shown in Figure 6.4. (Again, I think I'll leave those helpful data labels *off!)*

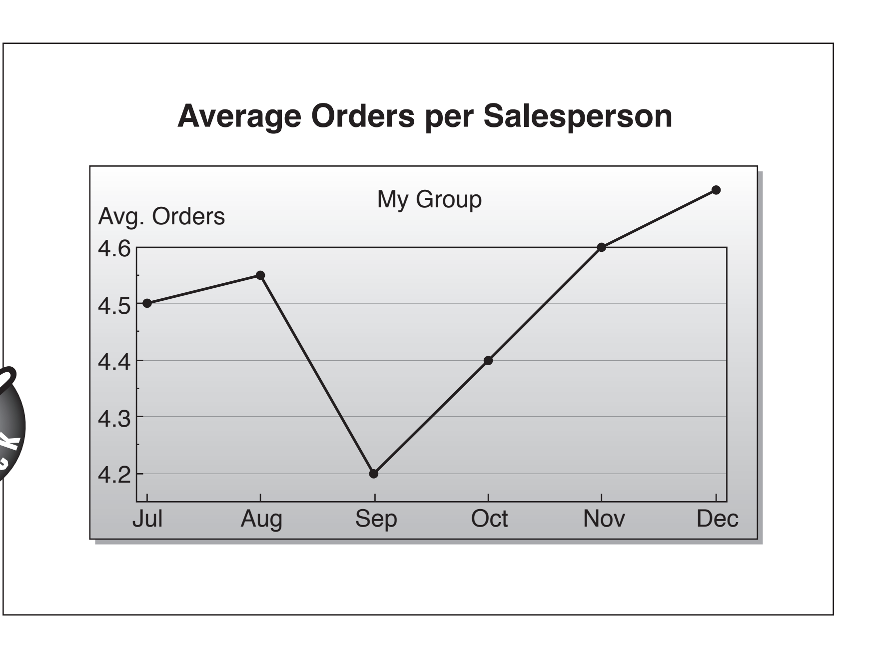


Figure 6.4 *Making the axis range narrower exaggerates fluctuations in the plot, which suits my deceitful purposes. I'm also pleased that the high point of the plot extends beyond the maximum value on the scale, which unnecessarily exaggerates the magnitude of that data point.*

What an improvement (at least, from my deceitful point of view)! But, notice something else about my clever chart: Expanding the scale emphasized the valley, as well as the peak. True, it now looks as though my team is making a big comeback, but the audience might also wonder why we took that big plunge in September.



To emphasize a plot—maximizing its fluctuations—*decrease* the *y*-axis range.



What, you might well ask, is the "honest" way of plotting the data? The most truthful version of any chart will be the one that has the least deliberate distortion. Figure 6.1 isn't bad, because its *y*-axis range is a typical one for the data being shown. But to be entirely fair to your audience—you should show that the *y*-axis scale doesn't start at zero, as described below in "Fighting Fair with Unconventional Axes."

Real Movers Don't Start at Zero

$$$$INVESTOR'S TIP$$$$

Almost all stock market charts will have nonzero starting points. For any particular stock, the audience is used to seeing its prices within a relatively narrow range, which sets the x axis boundaries.

I may be a shameless liar, but I'm no fool—I don't want to get caught. I started the *y*-axis of the first chart at zero, the *y*-axis of the second at 4.4. The scales aren't consistent—they have different starting points and span different ranges. If I show the two charts one right after the other, some bright observer might spot my trick.

To hide my intentions, I'll try yet another trick. Here's what I'm thinking: *Unless I take pains to show them otherwise, the audience will usually assume that axis scales begin at zero.* Therefore, I won't label the nonzero starting point on the scale.

Wait! I've got an even better idea. Why don't I omit all but the top label on the *y*-axis scale? If the audience thinks that the scale starts from zero, I will have given the impression that the *rate of increase* in the sales plot is even bigger, spanning the range 0 to 4.7 instead of 4.2 to 4.7. The result is shown in Figure 6.5.



Liars won't label the starting point of the *y*-axis scale. The audience will have to guess where the scale begins, provided that they are attentive enough to even try. The wider the range and the more divisions it has, the easier it will be to get away with your lie. For example, if the axis goes from 200 to 800, label only the top division at 800. The audience probably will mistakenly assume that the scale goes from 0 to 800.

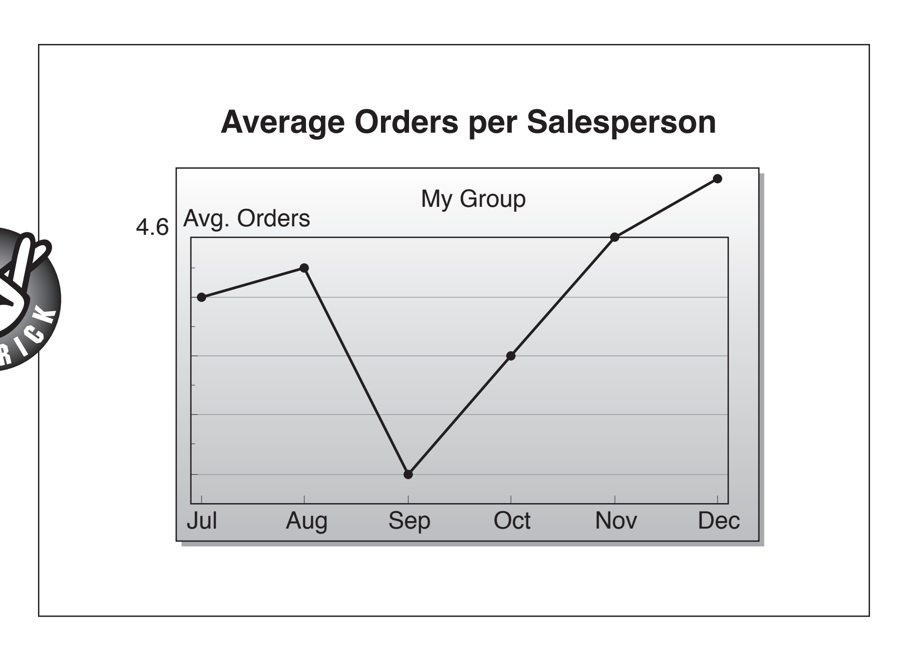
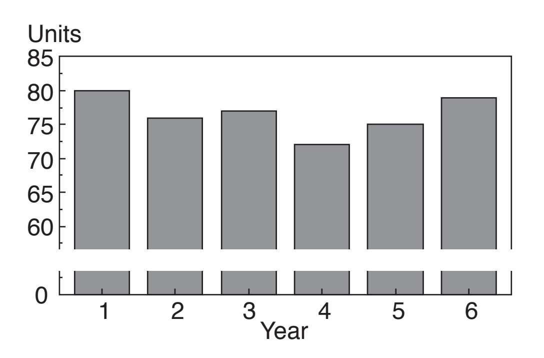


Figure 6.5 *Omitting most of the y-axis labels achieves a minimalist graphic design that is nonetheless grossly misleading.*

Fighting Fair with Unconventional Axes

Even if you have some legitimate reason for not starting an axis scale at zero, there is a way to fight fair so that the audience can see you're not up to any tricks. Start the scale at zero, then show a break in the axis line before continuing the scale. And if you want to be unmistakably clear, run the break across the full width of the chart:



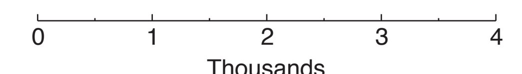
And don't make the scale labels too small!

Advanced Scaling for Ambitious Climbers

Scaling is another technique for manipulating axes. It involves keeping the range constant but multiplying or dividing its values by a number, or *scale factor.* For example, here's a scale that ranges from 0 to 4,000:

0 1,000 2,000 3,000 4,000

A less cluttered, cleaner-looking scale will result if you divide all its values by 1,000:



But notice that, despite the ranges of the two scales being exactly the same, the single-digit values on the second scale tend to deemphasize the magnitude of the data.



Closely related to axis scaling is the whole issue of labeling (or mislabeling, as liars tend to do), which is a major topic in Chapter 12.

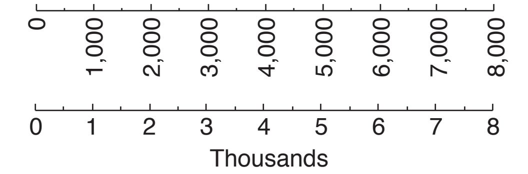
This trick of manipulation is kept honest by the scale label—Thousands. However, the trick appears more deceitful when the label is positioned farther away from the axis and perhaps also reduced in size, as in a less obvious footnote:

0 1 2 3 4

Units*

*Shown in thousands

The magnitude of the data can also be either emphasized or minimized by increasing or decreasing the number of scale divisions and their corresponding labels. For example, of these two scales, the axis on top emphasizes the scale magnitude more—even though their ranges are exactly the same:



The trick of increasing or decreasing the number of scale divisions can also affect how the audience perceives the *precision* of your data. For example, a scientist who is presenting experimental data could give the (perhaps misleading) impression of precision by including both major divisions and very fine minor divisions:

0.00 1.00 2.00 3.00 4.00

In reality, the data points might only be accurate within a fairly broad range—say, plus or minus one whole number (±1.0). However, dividing the scale into tenths would imply that the accuracy of the data points is much better—say, ±0.1.

For Extra Credit: Dueling with Dual Y-Axes

There are occasions—legitimate ones—for showing plots of two widely different magnitudes on the same chart. For example, if my company's labor costs are on the increase, I might want to show that curve as compared to U.S. Social Security receipts—a big factor in labor costs. My objective is to show that, although our costs are increasing, some huge trend that is outside my control might be pushing them up (so it's not really my fault!).

The basic problem, though, is that my company's sales are in millions of dollars, and the country's GNP is measured in trillions.

Plots of different magnitudes can be shown on the same chart by using two *y-axes*, with one scale on the left and a different scale on the right, as shown in Figure 6.6.

There's nothing inherently dishonest about such dual-*y* plotting. But as with the other techniques I've shown you, this one has its abuses.

Of course, there *is* a causal relationship between the expenses of maintaining the Social Security system and a company's labor costs. However, for purposes of my presentation, matters of national social policy are probably not at issue. My managers want to know what I

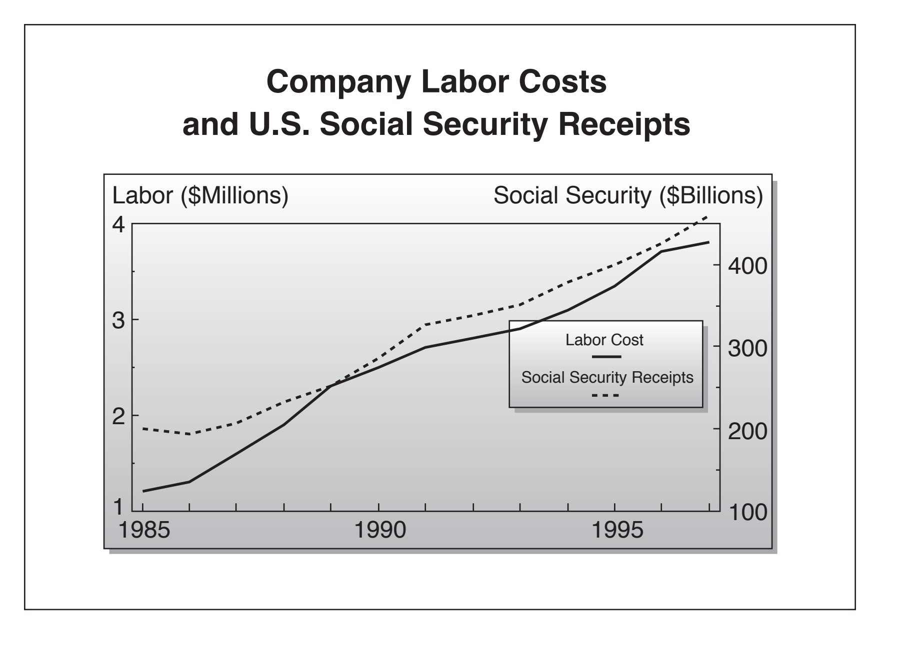


Figure 6.6 *A chart with dual* y-axes *can be used to compare plots of widely different magnitudes.*

$$$$INVESTOR'S TIP$$$$

The typical stock market chart has dual *y* axes, one for stock price and the other for trading volume.

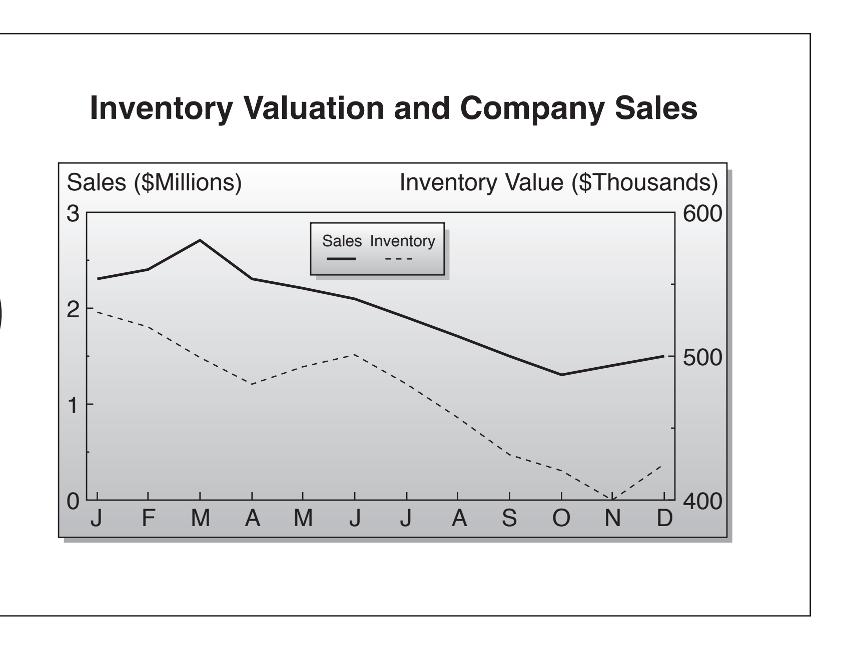
can do over the short term to control costs. Even though the national economic trend probably has some effect on my situation, such large issues just aren't relevant to my presentation.

To be both clear and honest in a dual-*y* chart, show just two data series—one for each *y*-axis. As a further help to your audience, show a legend or color-code the plots to their respective axes. For example, use yellow for one plot and its left *y*-axis scale, green for the other plot and its scale.

A clever liar might also use dual-*y* plotting to make bogus comparisons. Look at the chart in Figure 6.7. Sales volume is plotted against the left vertical axis, the dollar-value of inventory on the other. The implied comparison is that the sales results are affected by the availability of product. This could be a valid comparison, although the truth is probably the other way around—that inventory is restocked depending on how fast the products are selling.

It's also possible that the lying chart maker knows perfectly well that the decline in sales has nothing to do with inventory. The liar might be trying to hide much more significant causes, such as cutbacks in sales staff or advertising.

Figure 6.7 *It's not necessarily incorrect to show dissimilar data sets on the same dual-y chart. But is the comparison valid?*



The Big Number Crunchers Fight with Logs

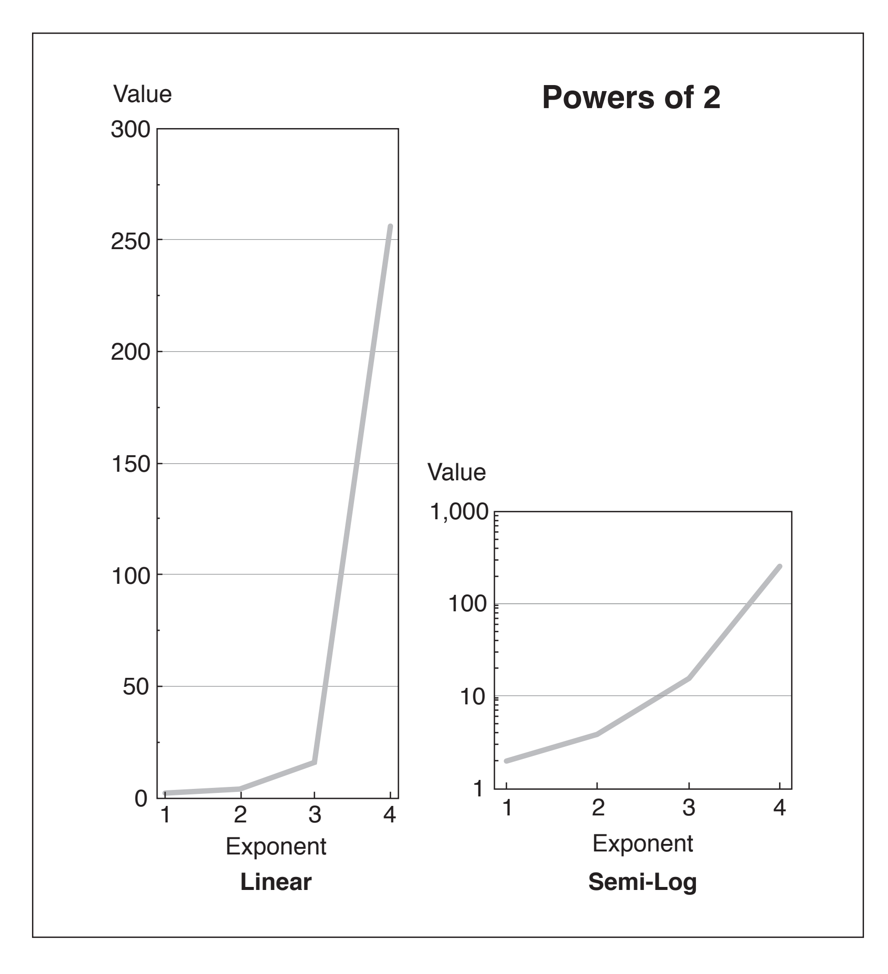
One of the more unusual tricks you can try with scales is to make them *logarithmic*. Logarithmic axes, or *log* axes, are scaled by powers of 10 (also called *orders of magnitude*). The legitimate purpose is to be able to fit a very wide-ranging data set on a single, compact chart. There really is no practical alternative to log scales—for example, if the points in the same data set range from the tens to the millions.

Although log scales in *xy* charts have their uses, this type of charting will not be easily understood by general audiences. Logarithmic plotting is used mainly by scientists and engineers, who typically deal with wide-ranging data, such as astronomical distances.

The most common (and most readily understood by nontechnical audiences) type of log plotting is called *semi-logarithmic,* or *semi-log*. The scale of only one axis—usually the *y*-axis—is logarithmic. The other—the *x*—is normal, or *linear*, as shown in Figure 6.8.

Both charts in the Figure plot the same data, the powers of the number 2: $2^0=1$, $2^1=2$, $2^2=4$, and so on. The chart on the left is the normal kind—

Figure 6.8 *A semi-log chart can shrink a wide-ranging data set down to size, but this plotting style might be unfamiliar to general audiences.*



$$$$INVESTOR'S TIP$$$$

Scientists and engineers are just about the only audience that won't be thrown by log charts.

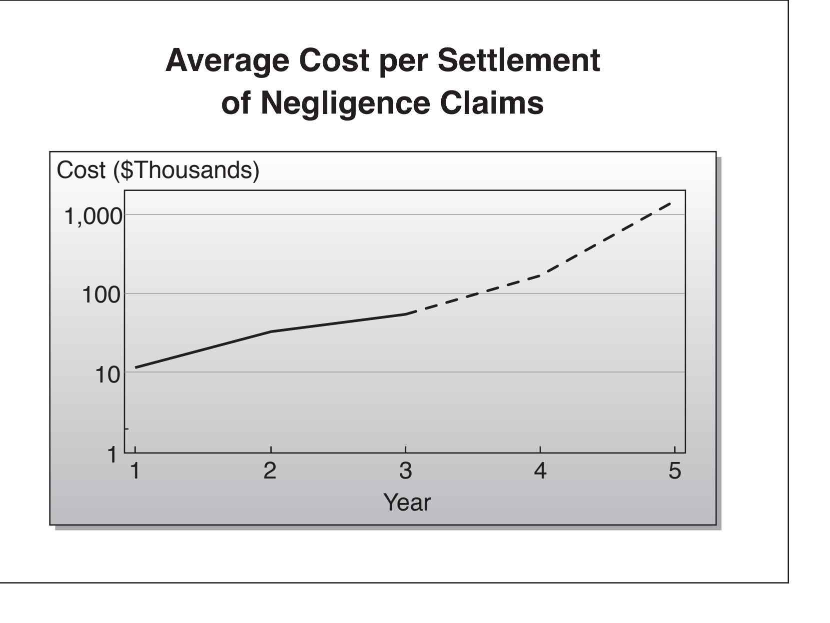
called *linear* because none of its axes is logarithmic. The one on the right is semi-log: The *y*-axis is logarithmic and the *x*-axis is linear. (More about a third type, log-log, shortly.)

Notice in Figure 6.8 that the spacing of the horizontal gridlines on both charts looks normal—until you check the *y*-axis labels. Here's another opportunity for mischief: The spacing of the grids can *appear* linear. Unless they study the scale labels, an audience won't guess right away that the chart is logarithmic!

Now, why would a liar want to use log scales?

In mathematics, wide-ranging trends that increase by powers are called *exponential*. Plotted against conventional linear scales, an exponential

Figure 6.9 *An exponential trend can be made to look less dramatic by plotting it on a semi-log chart, where it resembles a linear trend.*



curve can scoot upward at a shocking rate, as shown on the chart on the left in Figure 6.8. However, plot the same curve on a semi-log scale, as shown in the chart on the right, and it begins to look more like a straight line, which can be much less worrisome. An audience unskilled in interpreting log charts might well be led astray!

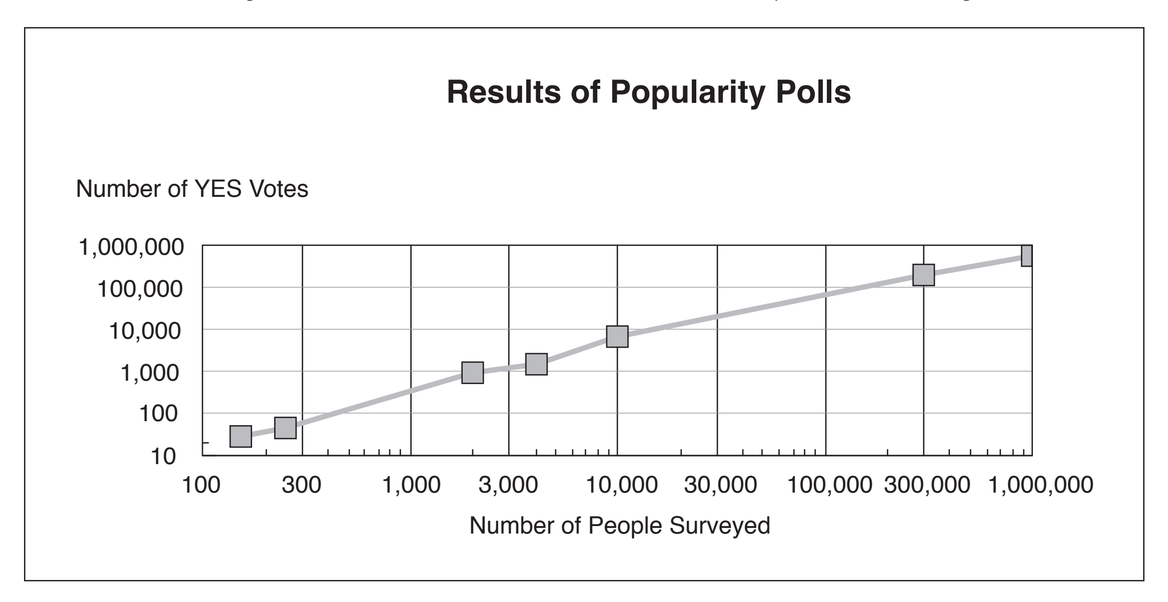
There's a practical example in Figure 6.9. Here the presenter is attempting to make an alarming increase in costs look like a less dramatic linear trend.



It's a cliché of modern media coverage that "exponential" costs/crime/environmental impacts are "threatening to spiral out of control." Despite the facts that the actual trends might not be truly exponential or that exponential trends don't spiral, in popular usage the term *exponential* often conveys the notion of looming crisis. For people who might otherwise worry about these things, there's the false comfort of the semi-log scale!

As shown in Figure 6.10, a doubly obscure charting type called *log-log* uses log scales for *both* axes, *x* and *y*. Don't resort to this technique unless you are sure that your audience is used to seeing data plotted this way. But if you do get out the double logs, be sure to use gridlines on both scales—and label them clearly!

Figure 6.10 *Log-log plotting can turn curves into straight lines that are easier to interpret, as in compound-interest and common-growth charts. However, nontechnical audiences may find it bewildering.*



Your Mission, If You Choose to Accept It

It should be apparent from this brief survey of potentially lying tricks that you can make an *xy* chart say just about anything by adjusting its axis scales. If you had any preconceived notions that charting is an objective science, these should be dispelled by now! Ultimately—and this should come as no surprise—you have to decide whether a chart you have made correctly conveys your interpretation of the data. There's no pretension to *truth* here. Charting is a medium of communication, and I simply want you to communicate what you intend. And I'm not so much worried that your intention is to lie through your charts. The real tragedy—for everyone involved—would be to lie without knowing it.

This discussion of axes applies to all *xy* chart types, regardless of the style you use to plot the data—whether lines, bars, or areas. In the next chapter, I'll describe how your choice of plotting style also affects the interpretation of a chart.

7

Chart Picks and Liars' Tricks

Or, Choosing Chart Types and Styles

$$$$INVESTOR'S TIP$$$$

It would be unusual, but not unheard of, for a pitchman to apply these tricks to deliberately mislead investors. Fudging in courtroom presentations has also been alleged. More often, the mistakes will be unintentional, but the results equally unacceptable.

IN previous chapters, I discuss various chart types and how they can be used—or abused. Here's some advice on picking the right charts for your presentation. There are also lots of pointers on knowing when someone else has deliberately picked the wrong chart for the job!

Selecting a charting format shouldn't be mainly about attractiveness of design or even personal taste. What you want to say to an audience about the data should determine the type of chart you choose to make your point. For example, you've probably grown weary of my preaching that pies are for percentages. But if that's your message, you should stick with pies. Also be aware that, having selected a chart type, important parts of your message will then be out of your control: The type of chart you select will make a specific statement about the data—regardless of how you adjust the design to suit your message. There are messages inherent in the chart types themselves, just as there are implied meanings in graphic arrows that point left, right, up, or down.



If you don't believe me about the nuances of left/right/up/down, you might want to review my words of caution in Chapter 3.

So, making these choices isn't necessarily about picking what's pretty. It's more like grabbing the right tool for the task. This chapter should give you some guidelines so that what they think you mean comes as close as possible to what you think you told them!

Being Picky about Types and Styles

For the sake of discussion, let's distinguish between *chart types* and *plotting styles*. In this book, there are only three major chart types: pie, *xy*, and some special ones. One special type already covered is the radar chart, which has its own unique set of rules. I'll mention a few other special types in this chapter, but it's mostly about the trouble you can get into with *x* and *y*.

Plotting styles are subtler variations within a chart type. *Three-dimensional* and *exploded-slice* are two styles within the pie chart type. Sometimes, styles can be mixed within the same chart. It's perfectly okay, for example, to create a 3D pie that also has an exploded slice.

Selecting a charting method involves picking a chart type, then plotting styles within that type. *XY* charts give you by far the widest range of choices. This major type includes several different plotting styles—which, if you want to be truly picky, might be called *subtypes* rather than styles. Subtypes would include various kinds of bars, lines, and areas used to define and connect a series of data points. And there can be lots of variation within each of those.



Liars will choose chart types and styles that hide or minimize messages that might otherwise be obvious in the data. In other words, liars will deliberately pick the *wrong* tool for the job.

Having agreed on these distinctions between type and style (okay?), let's see how your choosing among them can affect the conclusions an audience might draw from a chart's data.

Let's start by trying to plot the same set of data in a variety of ways. (This plan might not always work or even be a good idea, but let's see where it leads.) Table 7.1 shows the data in the raw, a seemingly unremarkable set of statistics, subject to no particularly alarming interpretations. It summarizes the number of ounces of liquid I drank each day in a typical work week. (This is just an example, so don't worry too much about why I bothered to keep track.)

Look first at the data series, which includes two rows of values. The top row is the days of the week. In mathematical terms, this top

Table 7.1 *Weekday Consumption of Liquids (oz.)*

MON	TUE	WED	THU	FRI
39	42	47	60	73

row is an *ordinal* set of values. Ordinal values imply some order, or sequence. The numbers 1–9 are ordinal because I can count from 1 to 9. The days of the week are ordinal because I can "count" from Monday to Sunday.

By contrast, there is no order of counting for the amounts of liquid I consumed. The numbers appear to occur randomly, although there might be some other as yet undetected pattern to the data points besides their occurrence in a daily sequence.

In this example, the ordinal set of values—the days of the week—controls the order in which the other set of values can be plotted on a chart. Because the ordinal values happen to be names of things (days), they can also be used as labels to identify the data points.

Pie

That's just what I need to make a pie chart—a set of numbers and a corresponding set of labels. Each number will determine the size of a pie slice, each label a slice name. To calculate the size of a slice, divide the number by the total of all the slices, then multiply by 360 (the total number of degrees in a circle). The result will be the size of the slice in degrees:

$$\frac{\text{NUMBER}}{\text{TOTAL}} \times 360 = \text{SLICE IN DEGREES}$$

After doing this for all the numbers, I have a chart (Figure 7.1) that shows the *relative amounts* of my consumption each day.

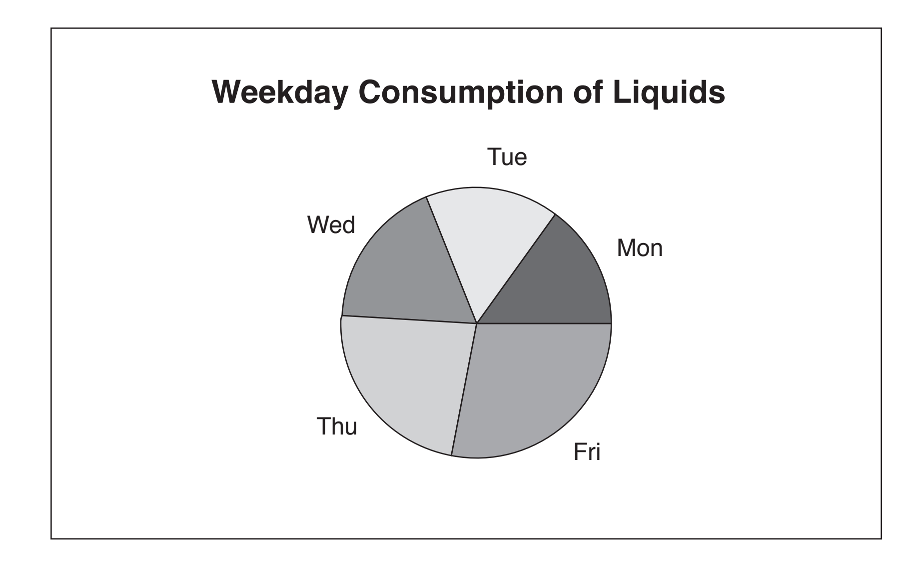


Figure 7.1 *A pie would be a way of comparing the relative amounts I consumed each day.*

The only use I could have for such a chart would be for making comparisons among the days' results. By looking at the chart, I could draw conclusions such as, "I drank a lot more on Friday," or, "I drank almost twice as much on Friday as I did on Monday."

You shouldn't be asking questions about specific amounts on this chart, for reasons explained in Chapter 2. (You shouldn't be thinking about actual values when you are looking at a chart that shows proportions.) Nor should I, the scrupulous presenter, be telling you about the amounts as you look at the chart—and you shouldn't be inferring anything about them. Remember, from a pie chart, there is no way of graphically showing the specific amounts represented by the total pie or any of its slices. (I could label the slices with amounts, but the labels aren't features of the pie itself.)



If I showed you a pie chart, but talked about the specific amounts of the total pie or of the slices, I'd be drawing the wrong conclusions. I might be simply misguided. Or I might be lying, because you have no way of knowing these things from the chart. You might be impressed with my *discussion* of the total, which is completely unsupported by the chart. Worse, you may remember this favorable impression long after you've forgotten what I actually said.

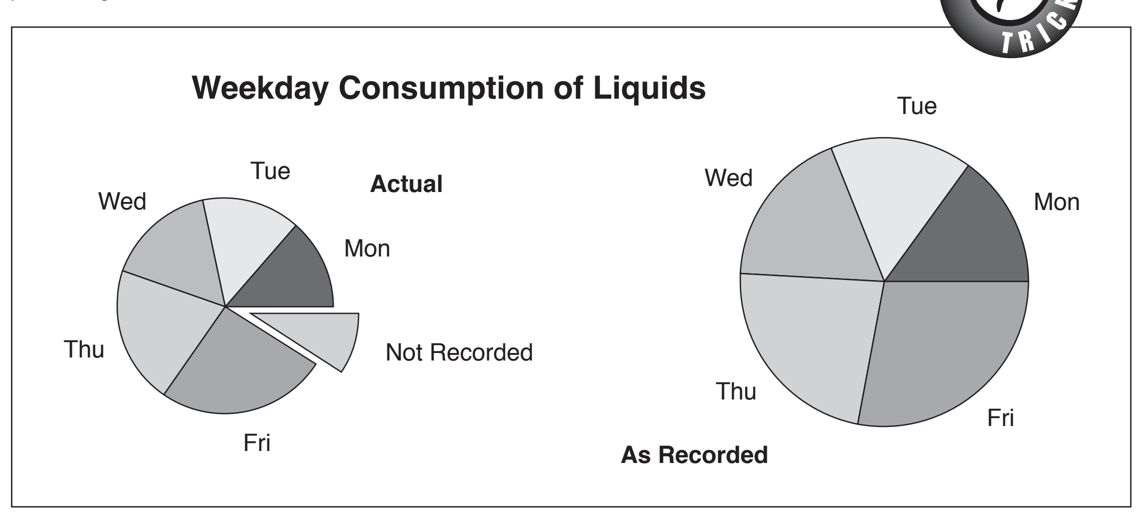
The fact that you can't know anything about the absolute total of a pie opens the door to further trickery. A liar might advise, "Don't show the 'don't-knows.'" Because the audience has no knowledge of the total, you aren't forced to show all the data. No matter how many slices you include, they will always total 100 percent of the pie. Omitting the "don't knows" means that, rather than plotting a slice labeled All Others or Don't Know, you simply don't include it. The effect will be to make all the other slices bigger than they would be if you had plotted all the data.

In my example, it could be that the data in Table 7.1 is only a partial record of my intake. I might have forgotten or neglected to make note of everything I drank. But when you see the chart in Figure 7.1, the implicit assumption is that I'm showing you 100 percent of the actual data.



A pie-chart liar's trick is to omit miscellaneous slices. If the miscellaneous items make up 10 percent of the actual data sample, the net effect will be to expand the remaining 90 percent to become 100 percent of the pie. In other words, the percentage of each of the slices will be exaggerated by making each of them 10 percent bigger. (See Figure 7.2.)

Figure 7.2 *The pie on the left includes all the data in an actual sample. The Not Recorded slice has been omitted from the pie on the right, thereby distorting the percentages of all the other slices.*



Line

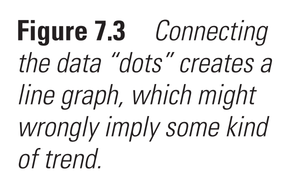
Okay, so you want me to talk about specific amounts. I'll have to resort to some subtype of *xy* chart. In perhaps the most common variation of *xy* chart, if one of the sets of values is ordinal, it must be used as the *x*-axis labels. The labels are placed in sequence at major scale divisions.

The other set—the numbers—becomes the *y*-axis values. If I plot the data points and connect them with line segments, the result is the line graph in Figure 7.3.

The advantage and the disadvantage of a line graph is that it highlights any trend in the data. In fact, line graphs are used so frequently to show trends that audiences are prone to see trends in those charts even if actual trends don't exist in the data.

Returning to the example, if you tried to find a trend in my drinking habits, you might suspect that I'd break my daily record on Saturday. You might also suspect that these anonymous liquids include alcoholic beverages and that my High Friday is just that!

For the moment, let's assume that you're right and that the data is a weekly trend and at least partially descriptive of my barroom habits.



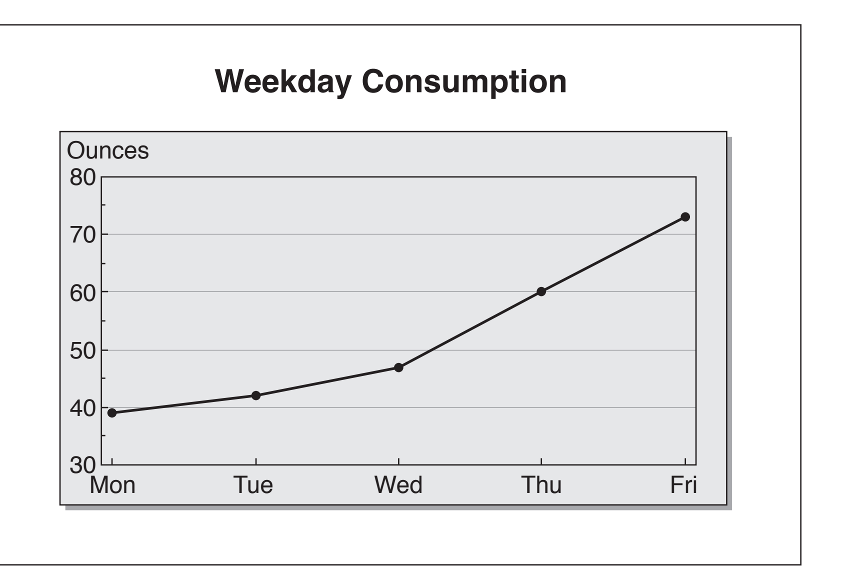


Figure 7.3 *Connecting the data "dots" creates a line graph, which might wrongly imply some kind of trend.*

But because I'm an accomplished liar, I can make you doubt that conclusion.

Recall in Chapter 6 that two dissimilar sets of data can be plotted on the same *xy* chart by adding a second *y* axis. I'll enhance the existing chart by adding a second *y* data set—the outdoor temperatures for the week in question. I then present you with the chart in Figure 7.4. What conclusion do you now draw?

It's not that I'm a problem drinker, after all. My intake was obviously done to replenish vital bodily fluids. You can see that the peaks in my drinking coincided roughly with the high-temperature days. Not guilty! It was lemonade, and I was in the midst of a heat wave! (Notice that I've also used the tips in Chapter 6 to adjust the *y*-axis scales to my advantage.)



Trendy thinking (about numbers) is the subject of the next chapter.

In reality, the outside temperature may have had no influence whatsoever on my drinking habits. In my first try at making a chart, perhaps I made a mistake by plotting it as a line, tempting you to draw inferences about a trend that might not exist. On the second try, I have both emphasized and explained that trend by superimposing another

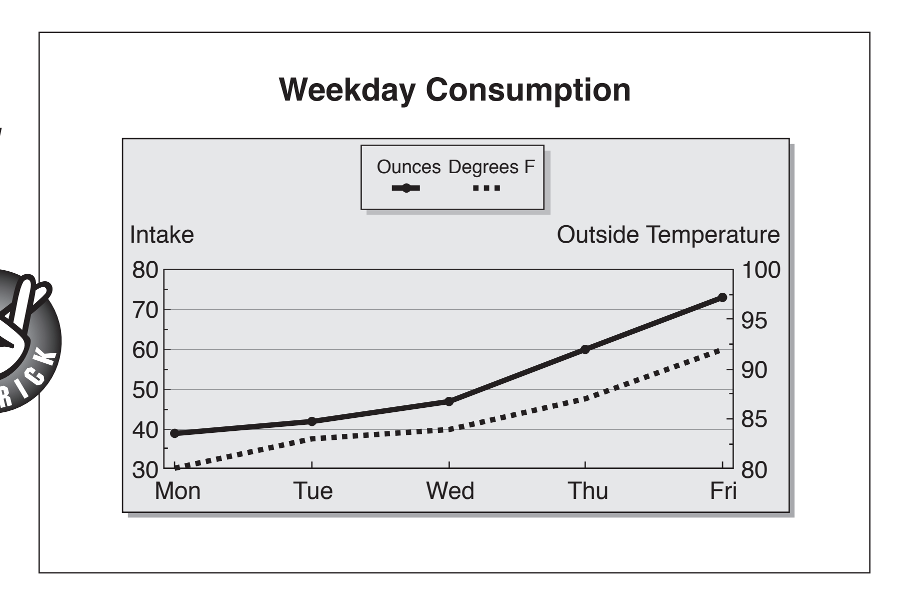


Figure 7.4 *Here's a really slick lie. I plotted something that really is a trend against a second y axis, implying that one set of data somehow influences the other.*

line—this one an undeniable trend. Everyone knows that the weather follows trends. And you will freely admit that the weather is beyond my personal control. Isn't my consumption of liquids, then, also out of my control?

Very neatly, I've used one trend to excuse the other, even though there may be no real cause-and-effect relationship between them.



If I look hard enough, I can find some trend in the environment that has the same pattern as my initial data plot. I can point to the value of the yen, the balance of trade payments, population growth, or a decline in purchasing power. Computer programmers even say there's a trend for trends that defy other explanations—*POM-dependency* (phase of the moon). A clever liar will superimpose such an indisputable trend on a questionable one to give the mistaken impression that the first has a controlling influence on the second.

Area

If I wanted to emphasize the *volume* of my consumption, I'd choose an area chart, which is much the same as a line graph, but with the area underneath the line shaded, or shown as a solid area (Figure 7.5).

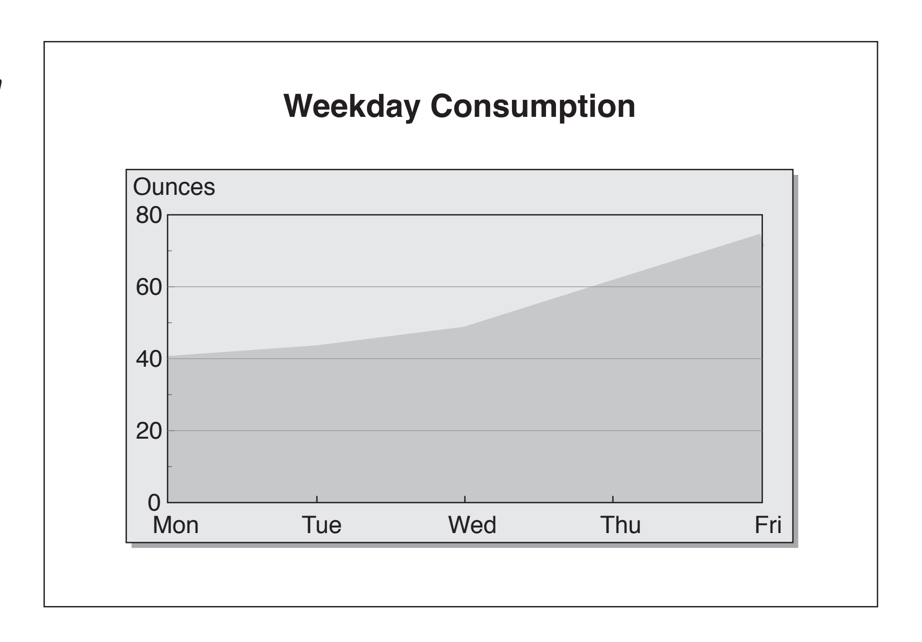


Figure 7.5 *An area chart emphasizes both a specific volume (*x *times* y*) and a trend.*

The total volume of my weekly consumption is 261 ounces, which is represented by the whole of the solid area. Unlike a pie slice, this area has a specific value, as measured against the chart scales. It can also show a trend, by virtue of the line graph that marks its upper boundary.



You can get a good approximation of the area by drawing the top of a rectangle through the midpoint of the line that defines the upper edge of the area. The size of the area is then $x \times y$, where y is the point at which the top edge of the rectangle intersects the y axis.

Perhaps the best use of an area chart is for comparing the specific volumes of several sets of data. Think of it as a hybrid of both pie and *xy* charts—showing ratios *and* specific amounts.

It's time I came clean. Table 7.2 breaks my weekly consumption of liquids (in ounces) down into different types of beverages.

Figure 7.6 plots this data as multiple areas on the same set of *xy* axes.

Now, there's something you should understand about this area chart. It's data series are *stacked*, or plotted on top of one another. The base line of the first (bottom) area is the *x* axis (*y*=0). But the base line of the second (middle) area is the top edge (the line graph) of the first. The base line of the third area is the top edge of the second.

Table 7.2 *Breakdown of Intake by Beverage Type (oz.)*

	MON	TUE	WED	THU	FRI
COFFEE	6	6	10	6	8
SOFT DRINKS	16	18	15	20	32
FRUIT JUICE	4	0	3	10	3
WATER	13	17	19	24	30
DAILY TOTAL	39	42	47	60	73

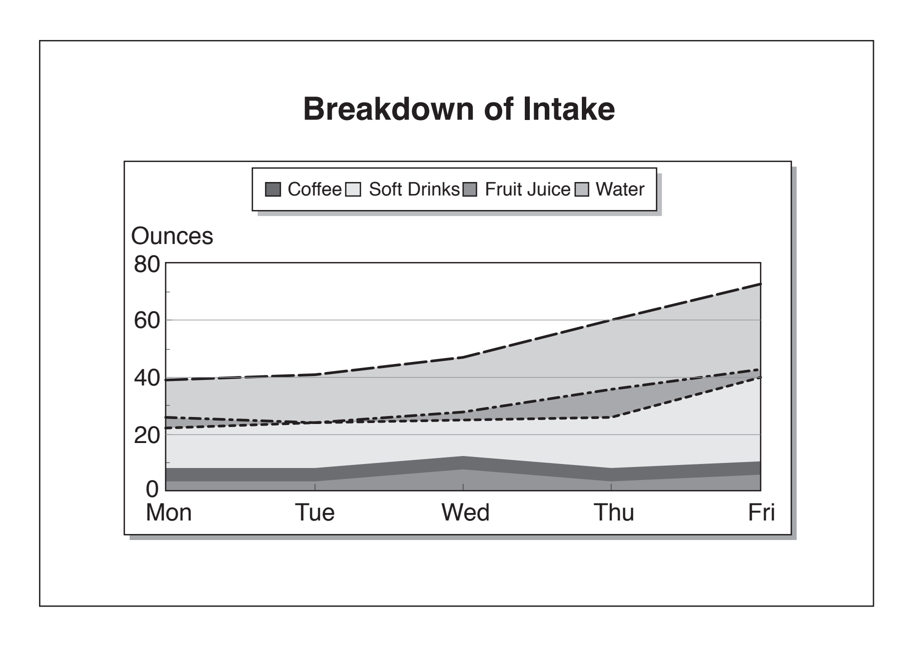


Figure 7.6 *Multiple areas permit the volumetric comparisons of pie charts, but with the specific values of an* xy *chart.*



Remember the rule for using stacked areas correctly: Put the data series with the least amount of fluctuation on the bottom. Otherwise, a "bumpy" area will create distorting bumps in the areas above it.

Your audience might not understand this stacked effect. Unless you tell them otherwise, some people will assume that the base lines of all areas begin at the *x* axis. Figure 7.7 shows the two ways of interpreting an area chart if the audience doesn't know whether—or understand how—the series are stacked.

An unambiguous way of showing different areas is to plot them in 3D, as shown in Figure 7.8. In such a chart, the areas are usually *not* stacked. Here it is important to show the area with the least values (not necessarily the least variation) in front, to avoid hiding the tops of the areas in back.

A problem with 3D areas, as with some other types of dimensional plotting styles, is that it can be difficult for the audience to estimate the exact tops of the areas in relation to the *y*-axis scale. However, since the purpose of an area chart is usually to emphasize volumes, this drawback might be less important than the potential problem of interpreting the base lines of 2D stacked areas.

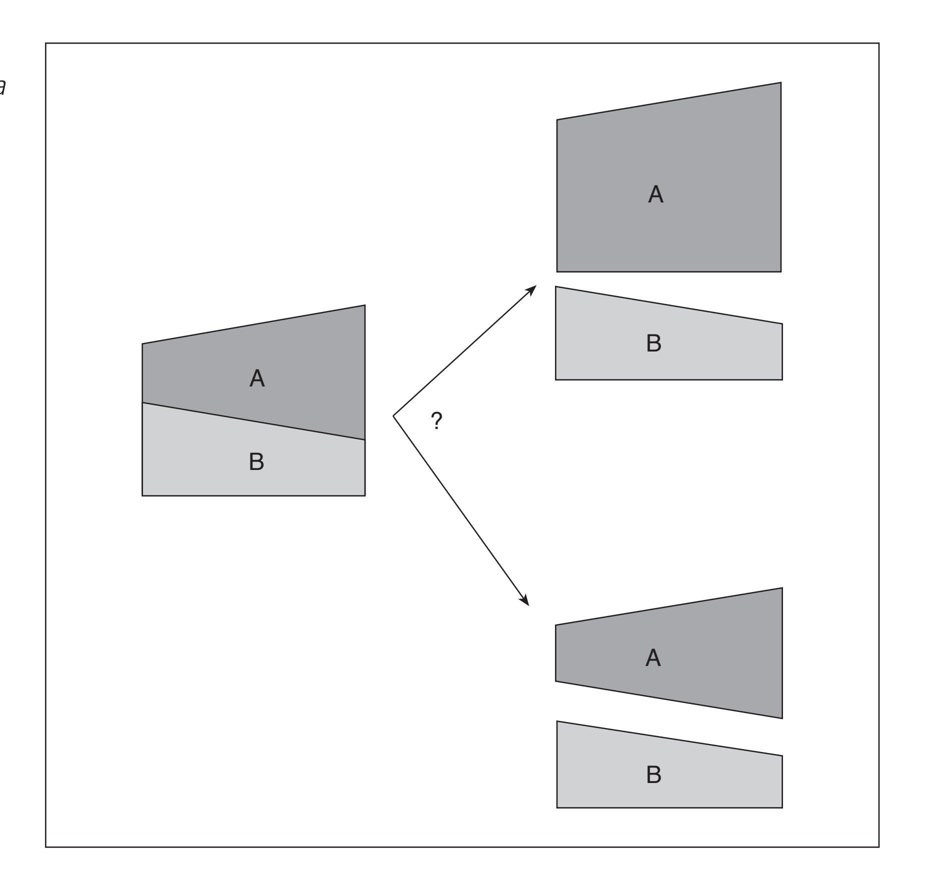


Figure 7.7 *If you don't know that the data series of an area chart are stacked, there are two ways to interpret how the areas are drawn.*

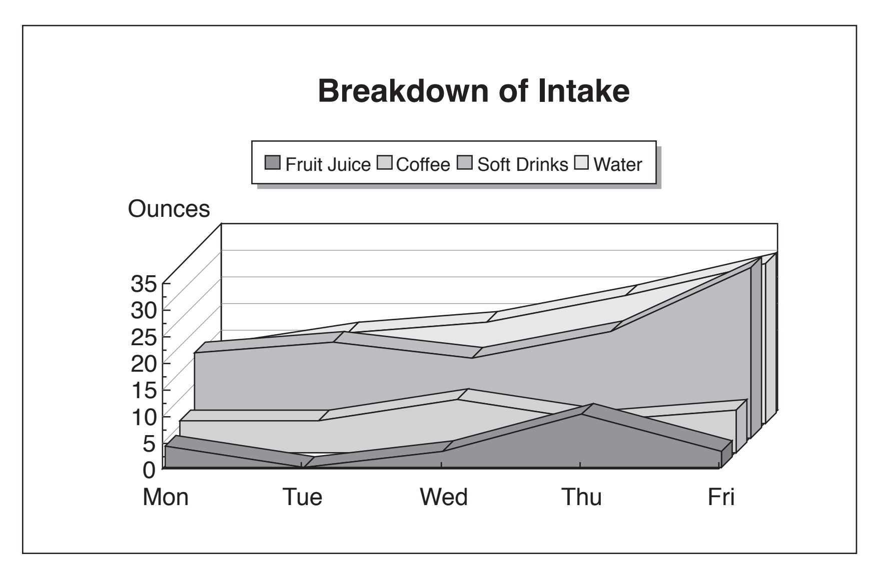


Figure 7.8 *A 3D area chart might be easier for the audience to interpret than the 2D stacked type. (Data series have been rearranged so that the areas are easier to see, but notice that dimensional areas are prone to the same type of distortion already discussed in relation to 3D bars.)*



The liar's trick with areas is a fairly simple deceit: Areas simply look more *substantial* than line plots. Because an area emphasizes total volume, the audience might be less concerned with the peaks and valleys than if the data were shown simply as a line.

Vertical Bar

Plotting the first set of data as vertical bars is likely to have a very different effect on an audience. Because bars are separate, they tend to minimize trends, emphasizing the individual result at each division of the *x* axis. The first set of data, describing my consumption of unnamed beverages, is shown as vertical bars in Figure 7.9.

If you insist on seeing trends in bar charts, there are two variations you can try. The first is to connect the data points at the tops of the bars with a line graph. This is called the *bar-line* style. (The line need not plot the same data series as the bars, but it does, in this case.) The second is to draw line segments between the top corners of the bars. This technique is called *linked bars.* Both techniques are shown in Figure 7.10.

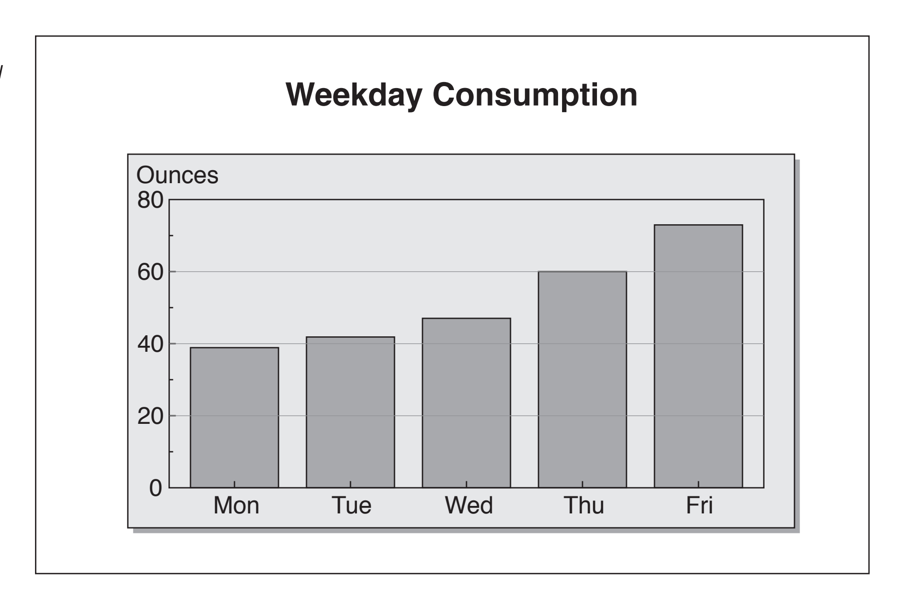


Figure 7.9 *Vertical bars minimize trends and emphasize the individual results at each* x-*axis division.*

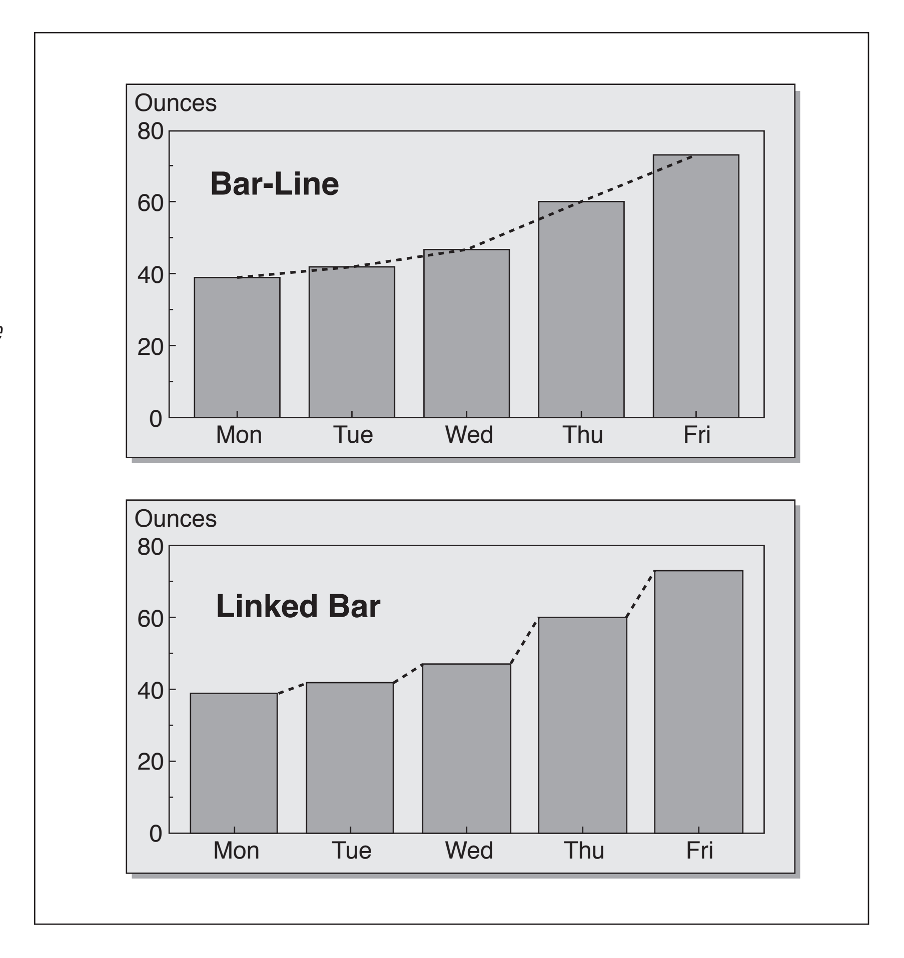


Figure 7.10 *Two techniques for highlighting trends in bar charts are bar-line and linked-bar styles. (Use of "fat" bars could be an unfortunate choice, perhaps implying some span of time, when, in fact, there is just a single data point for each day. Also perhaps misleading are the areas under the trend lines, which aren't meaningful in this type of chart.)*

Use these bar-line or linked-bar techniques with caution. A major reason to use vertical bars is to show results from multiple entities. The *x*-axis data need not be ordinal; the labels can refer to totally separate things. For example, a bar chart could show sales results, in dollars, by salesperson. The *x*-axis data, then, might be the names Jill, John, Jake, and Jane. Aside from alphabetical order, there is no obvious way to "count" from Jill to Jane. But—regardless of their order on the *x* axis scale—the labels still make a perfectly understandable chart (Figure 7.11). This is one of the few instances in *xy* charting when the *x* axis has nothing to do with the flow of time. Furthermore, results of

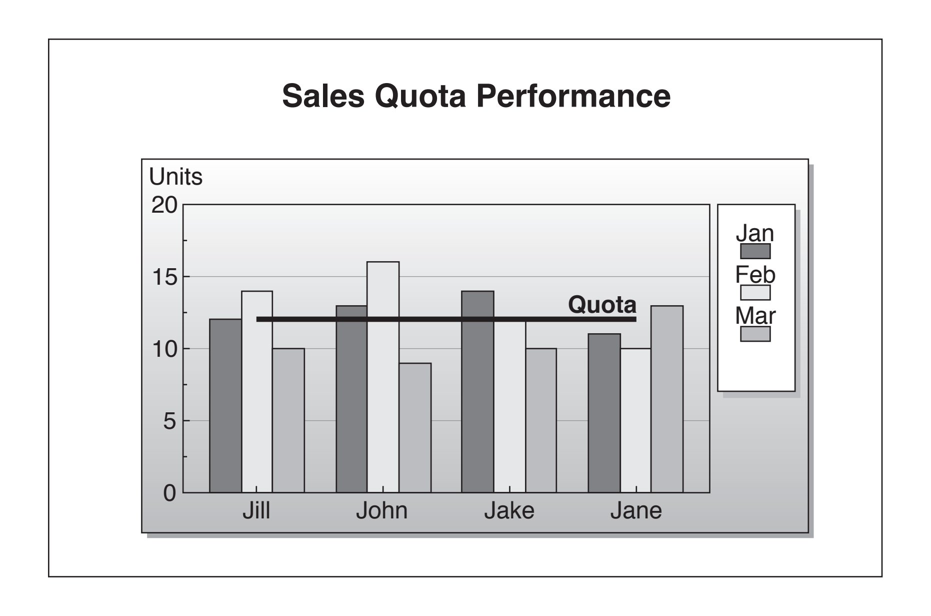


Figure 7.11 *The bar chart is the only type that can be used to show separate entities (salespersons, in this case) as* x*-axis labels.*

different entities can't be shown in any other *xy* chart style—you must use some type of bars.

If a chart must show multiple sets of compatible bars (all plotted against the same *y* scale), you have three different styles to choose from: clustered, stacked, and cumulative.

The data from Table 7.2, describing my consumption by day and type of beverage, is plotted in these different ways in Figures 7.12–7.14.

Clustered

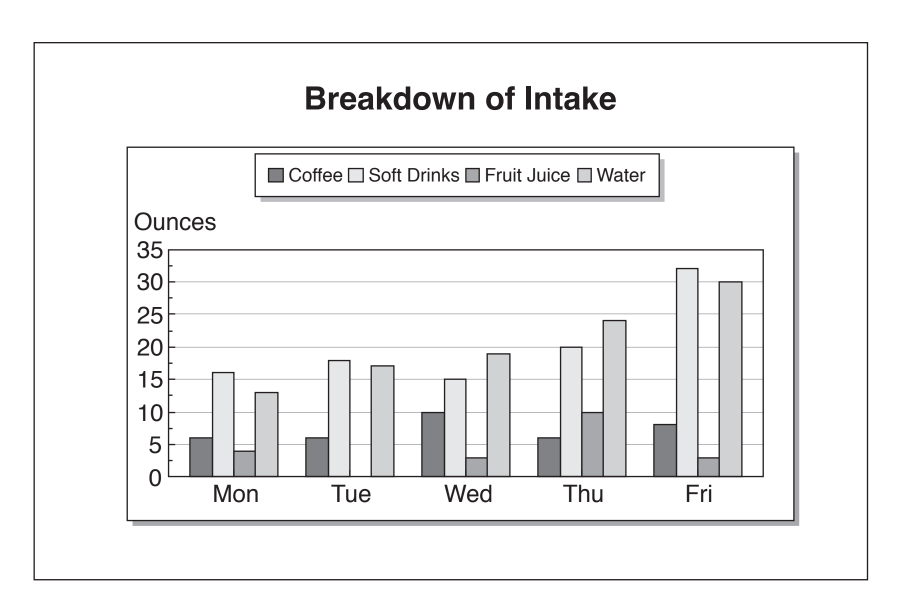
In a clustered bar chart (Figure 7.12) separate bars are plotted for each data series. One bar from each series is grouped with the others, forming a cluster, at each major division on the *x* axis.

If you use the clustered style for bars, it will be fairly easy for your audience to estimate the height, or amount, of each bar. The height is obvious because all bars begin at the *x* axis base line (y=0).

Stacked

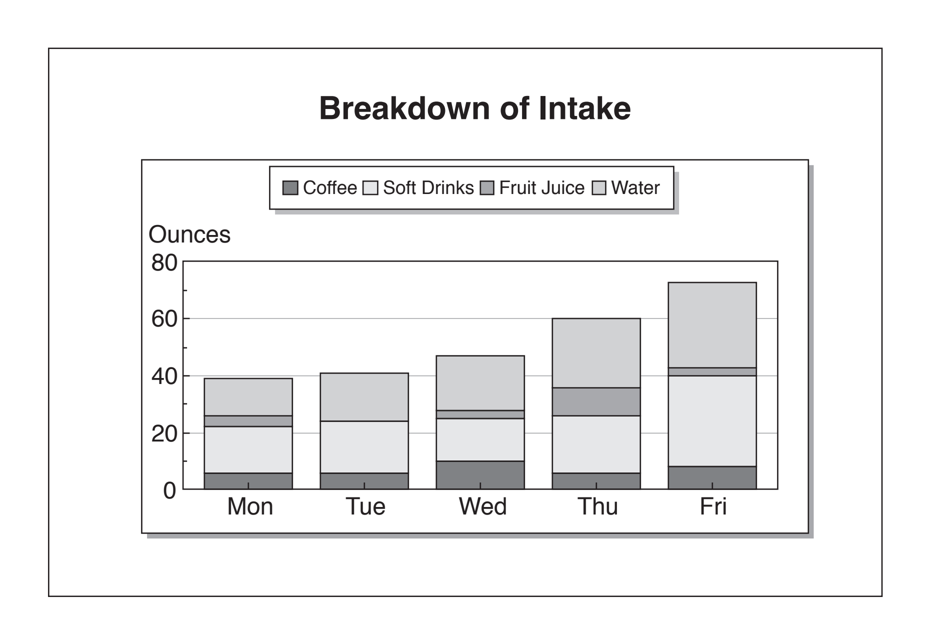
In the stacked bar style, bar segments for each data series are stacked like building blocks on top of one another at each *x* axis scale division

Figure 7.12
Clustered bars—useful for side-by-side comparisons in each time period.



to form a single bar (Figure 7.13). Compared with clustered bars, it will be more difficult for the audience to estimate the absolute sizes of each segment (since they *don't* all start at y=0), but it will be easier to compare the relative sizes of segments within each bar.

Figure 7.13
Stacked bars—notice that the total height of the stack is the sum of each bar's segments.



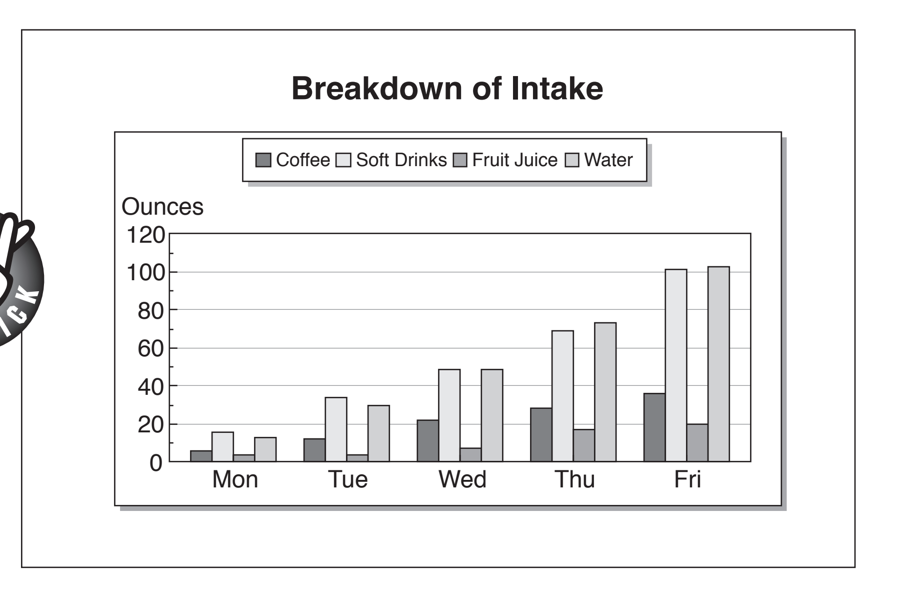
Cumulative

Bars can be *cumulative* only within a set, or data series (Figure 7.14). In a set of cumulative bars, the height of each bar represents a running total of the bars to that point. So, the value of the second data point is added to the first; and the result is plotted as the height of the second bar. The third point is added to the first two to produce the third bar, and so on to the end of the data set.

Cumulative plotting is fine for showing things that actually accumulate, such as the buyer's equity in a home or the compounded interest in an investment. Cumulative plotting is *not* appropriate—and can be a cheap trick—for amounts that aren't usually shown as running totals. Everybody's favorite numbers—sales results—would be an example. True, monthly sales accumulate during the year to produce gross revenue. However, the monthly results are usually not shown as cumulative because the audience wants to focus on how current or forecast performance compares to the results for the preceding months.

Now, here's a real note of caution: Many people confuse *stacked* and *cumulative*. Remember, *stacked* refers to multiple sets of data—

Figure 7.14 *Cumulative series, clustered bars. (Note the greatly expanded* y*-axis, which now spans 0–120.)*



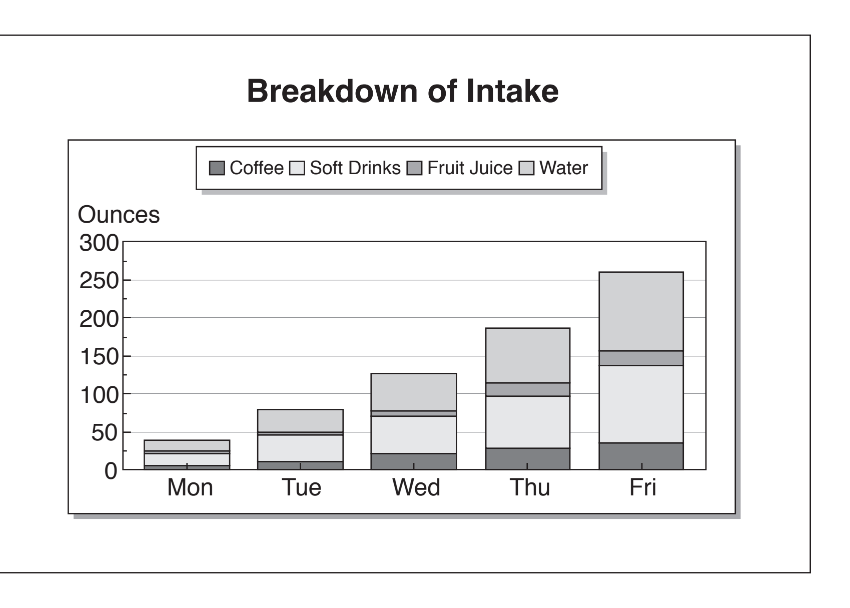


Figure 7.15 *The bars in this chart are both cumulative within each data series and stacked among series. The result is either hopeless confusion or meaningless optimism about those steadily climbing bars!*

one stacked on top of the other. In charting, cumulative plotting applies *within* just one set of data—as each point is added to the preceding ones.



It's a particularly nasty liar's trick to mix stacked and cumulative techniques in the same chart, especially without making mention of the cumulative part. The example in Figure 7.15 takes unfair advantage of the potential confusion of stacked and cumulative plots, showing bars that are not only deceptively tall but also steadily increasing in height—even though the noncumulative data is much less impressive.

100 Percent

A curious variation of the bar chart is *100-percent bars*. All bars have the same height, representing 100 percent of the thing the bar represents. Each bar, then, is much like its own pie chart. A set of 100-percent bars can be used instead of a multiple-pie chart, with the advantage that it will be much easier for the audience to make comparisons among bars than among pies.

Use 100-percent bars for the same reason you'd use pies—to show percentages and ratios. Although multiple pies are usually confusing

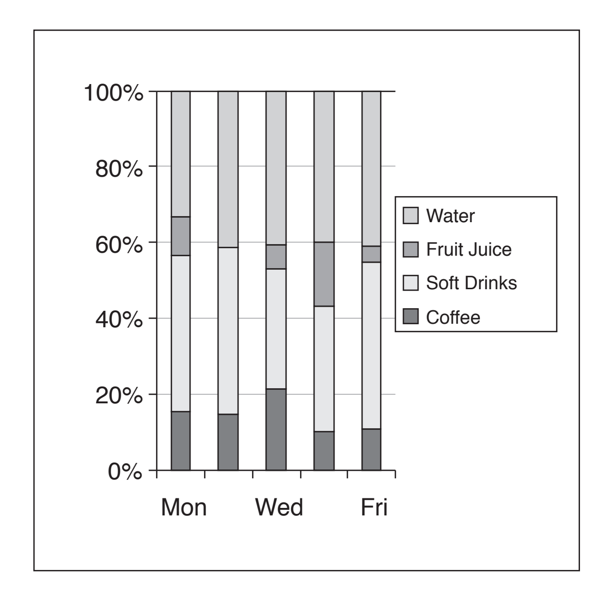


Figure 7.16 *The 100-percent bar style can be used instead of multiple pies to compare ratios.*

and difficult to read, the example of 100-percent bars in Figure 7.16 shows how the same objective can be achieved in a much more understandable chart.

Horizontal Bar and Gantt

To recap my Theory of X and Y discussed in Chapters 3 and 4, up-down means gain and loss, and left-right usually means the passage of time.



It's a liar's trick to show things that depend on magnitudes—such as sales—as horizontal bars. Save those lazy bars for showing time spans, such as the durations of tasks in a Gantt chart.

Paired Bars

A variation of horizontal bars—the *paired-bar* style—by its very nature permits you to transgress not one but two tenets of my Theory of X and Y. The misuse of this style shares the first violation with other horizontal

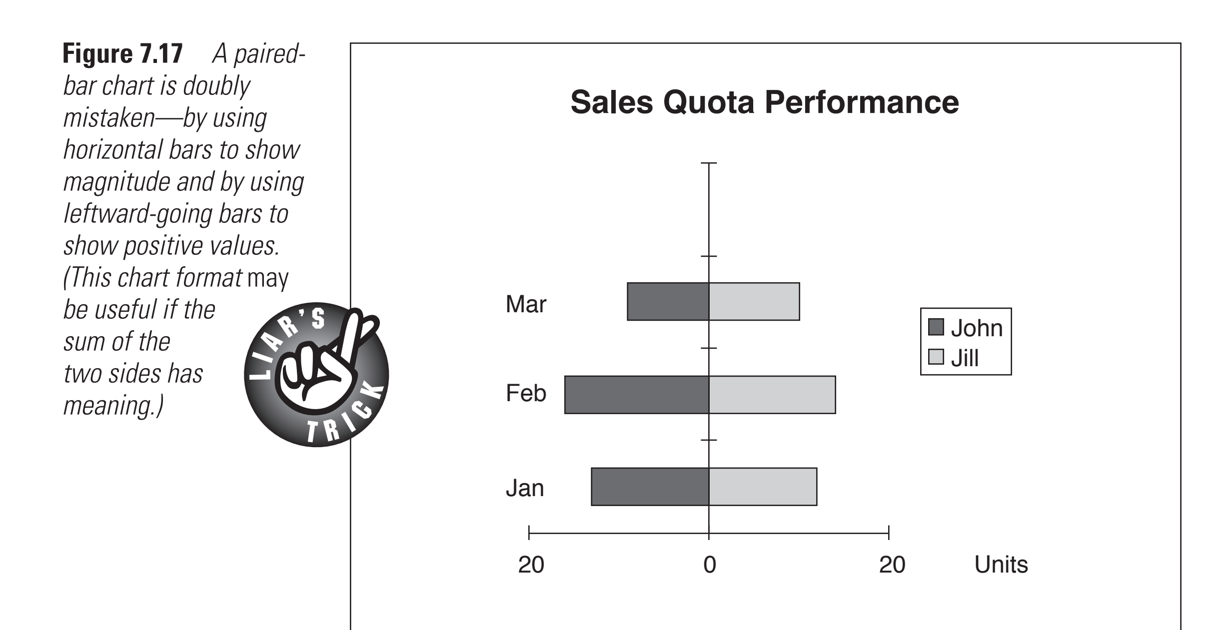


Figure 7.17 *A paired-bar chart is doubly mistaken—by using horizontal bars to show magnitude and by using leftward-going bars to show positive values. (This chart format* may *be useful if the sum of the two sides has meaning.)*

bar charts—showing magnitudes on a horizontal rather than a vertical scale. The second violation has to do with the pairing of the bars—one set grows rightward from an origin in the center, the other leftward. Now, recall that in the Cartesian coordinate system of plotting, a scale that proceeds leftward from the origin represents *negative* values. A paired bar chart, however, ignores this notion. The pairing of rightward- and leftward-going bars is simply for purposes of comparing the two data sets. So, in a paired bar chart, motion *outward*—to the right or to the left—means gain; motion inward means loss. There's an example in Figure 7.17, but, as you might expect, I'm not recommending this style to anyone.

Venn and Bubble

Circular kissing cousins in the chart world are *Venn diagrams* and *bubble charts.* If you remember your high school math teacher talking about *sets*, you're way ahead of the game here. In mathematics, a set is a group of things, usually numbers. Typically, some criteria or

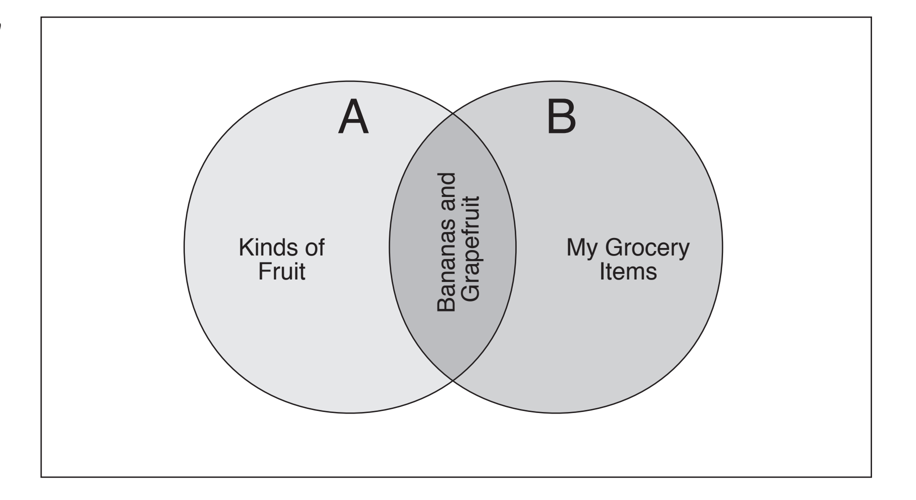


Figure 7.18 *This Venn diagram shows the intersection of sets A and B.*

characteristics define what stuff goes in a set, such as a set of "kinds of fruit." Another set might be "items in my grocery sack." The *intersection* of two sets is what happens when you compare their respective contents. The intersection of "kinds of fruit" and "items in my sack" turns out to be a bunch of bananas and a grapefruit. Figure 7.18 shows this in a Venn diagram as two intersecting, or overlapping circles.

The *union* of sets happens when the contents of two different sets are combined, as shown in Figure 7.19. In this case, the union of "apples" and "oranges" is "pieces of fruit," shown as two circles that are entirely contained within the larger set, or category. (A set that is entirely contained within another is called a *subset.*)



Venn diagrams are for showing concepts, not quantities. The sizes of the circles can relate to the relative sizes of the groups, but more often they must be sized to permit larger circles to contain smaller circles, showing relationships between subsets and the larger sets that contain them.

Bubble charts place Venn diagrams, in effect, on *xy* chart axes. The purpose is to show how sets of things compare according to factors such as complexity (usually the *y* axis) and time (usually the *x*), as shown in Figure 7.20.

Figure 7.19 *This Venn diagram shows the union of sets A and B.*

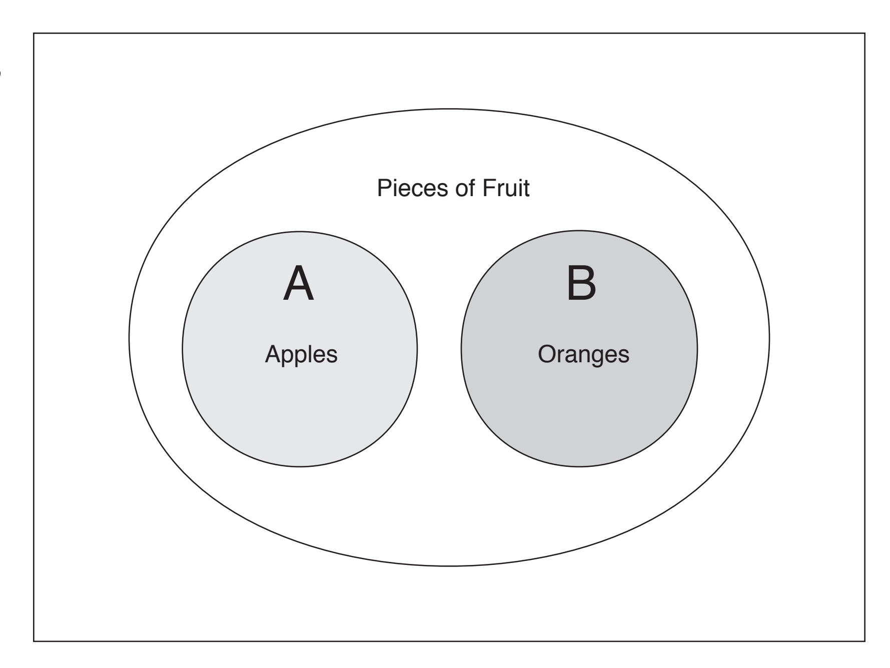
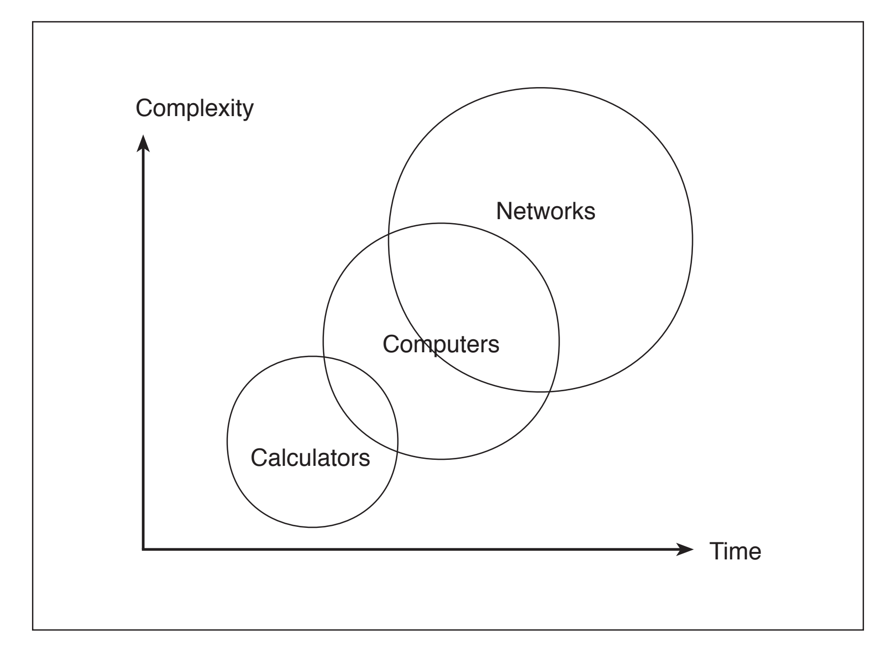


Figure 7.20 *Like Venn diagrams, bubble charts typically show concepts, not quantities. However, chemists use them for quantitative phase diagrams, and marketing managers place different products or product lines along price and performance axes.*



Despite this plotting scheme that looks suspiciously like an *xy* chart, bubble charts should not be used to show absolute quantities. The scales are relative, and so must be the sizes of the circles.

Mixed charts

If you try to mix several styles for different plots in the same chart, you should know that there are only a few combinations that are effective.

Refer back to Figure 7.10 to see a combination of bars with a line (a *bar-line* chart). Figure 7.21 shows points on a scatter chart with one of several possible trends traced through them as a line.

This topic of spotting trends in charts is so full of opportunities to do mischief that I devote the next chapter to pointing them out.

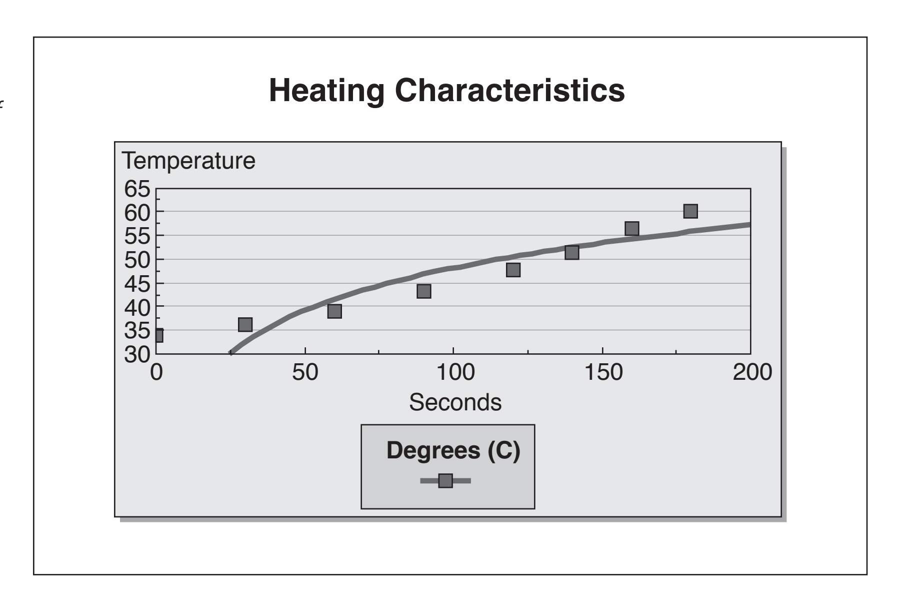


Figure 7.21 *Divining trends from discrete points is a primary use of the scatter-line chart.*

8

Trends

Or, A Conservative's Guide to Trendy Thinking

$$$$INVESTOR'S TIP$$$$

Spotting trends in markets is obviously a crucial focus of financial analysis. For example, the 65-day moving average in stock charts is used to highlight overall performance over short-term volatility. But remember that any such summary is a further abstraction of the data and is therefore dependent on the judgment of the chartmaker.

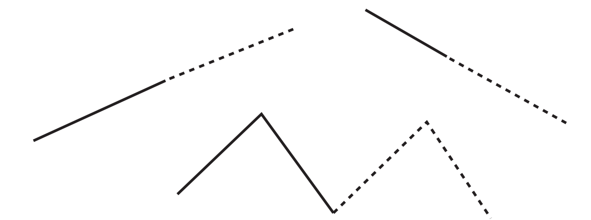
SINCE ancient times, seeing into the future has been an inexact science and an even more unreliable business. In Shakespeare's *Julius Caesar*, high priests dissect a dove to inspect its entrails, hoping to divine in that awful mess some clue about the Roman emperor's fate. Today, having mostly abandoned mysticism as a basic business tool, we look for signs of the times in charts. Ambitious business managers must now be enthusiastic trend spotters, using various *trend analysis* techniques to study squiggly lines as carefully as our superstitious ancestors examined those bird intestines.

Considering the theme of this book, you can probably guess that trend analysis, like other charting techniques, can be a tool for either enlightenment or deception.

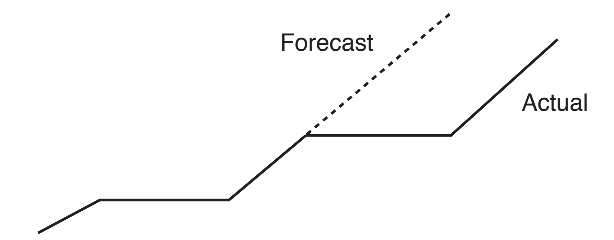
This chapter examines trends in *xy* line charts. Line charts can highlight trends in data better than any other chart type. In fact, they are used so often to show trends, that audiences are prone to view all line plots this way. They see an increasing *slope* (the techie term for line-chart *slant*) that they assume will continue upward, or a decreasing slope headed downward, or a cycle of ups and downs that must surely continue to repeat itself.



Some convention or other dictates that we show the present as a solid line, the tenuous future as dashed or dotted.



The ability to emphasize trends visually can be a powerful tool for the conscientious presenter, but it can also be just as effective when used by liars who have no qualms about exploiting our expectations. After all, a trend, if such a thing exists at all, might turn out to be something very different from the initial forecast:



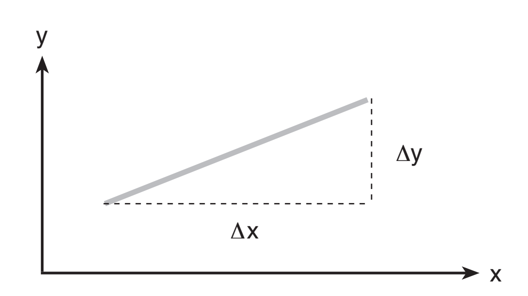
Perhaps to exploit this bias toward trendiness, line charts are commonly used by investment counselors. The intent is to show prospective

This Slope Isn't So Slippery

In an *xy* chart, *slope* has a very precise meaning. The slope of a line for any two points is the difference between the two *y* values (called *delta y* and written (Δy), divided by the difference between the two *x* values, or

$\Delta y / \Delta x$

If the line slopes upward, the second *y* value will be greater than the first, and the slope will be positive. If the line slopes downward, the second *y* value will be less than the first, and the slope will be negative. The higher the number, the steeper the slope.



investors how some pattern in the present—such as the return on an investment or the price of a stock—foreshadows future gain. However, as eager investors often learn too late, there is usually no assurance that the factors driving up today's income will continue to operate. So, the trend that looked like a trustworthy promise was, at best, an educated guess about what *might* happen.

Trend Analysis: Fact or Fiction?

Mathematicians approach trend analysis this way: Find a formula or equation that is capable of generating the data we can observe. The attempt is usually to find the simplest formula that successfully predicts future results based on numbers we have gathered in the past.

For example, faced with the data series:

0, 2, 4, 6, 8

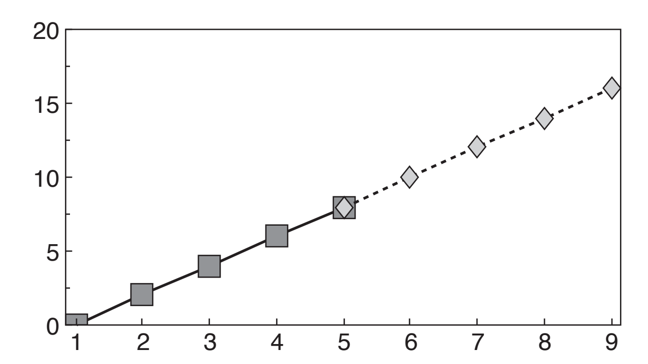
The mathematician might derive a formula something like:

NEXT NUMBER = PREVIOUS NUMBER + 2

Having found such a formula, a mathematician might then use it to predict future results:

0, 2, 4, 6, 8, **10, 12, 14, 16**

Viewed as a chart, the trend seems perfectly obvious.



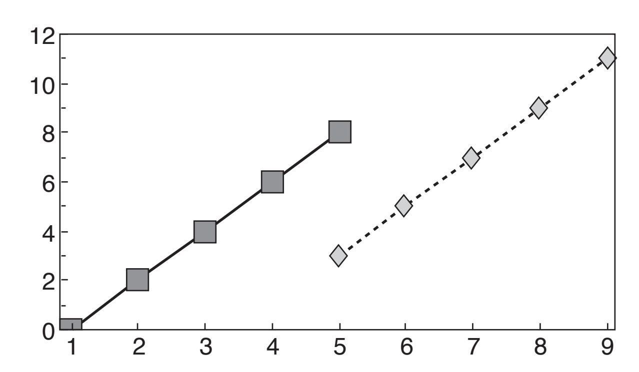
But the pattern in the actual data might not be so simple. When the future is now, so to speak, the actual data might look like this:

0, 2, 4, 6, 8, 3, 5, 7, 9, 11

There is a pattern here as well, but it isn't apparent from inspecting the first four numbers. The formula for this series might be

START AT 0
NEXT NUMBER = PREVIOUS NUMBER + 2
DO FOUR TIMES
START AT 3
NEXT NUMBER = PREVIOUS NUMBER + 2
DO FOUR TIMES

As with the straight sloping line in the example above, it's much easier to spot the trend in a chart:



The mathematician's formula amounts to a guess about the process that is generating the numbers. The pros have a word for extending the past into the future—*extrapolation*.

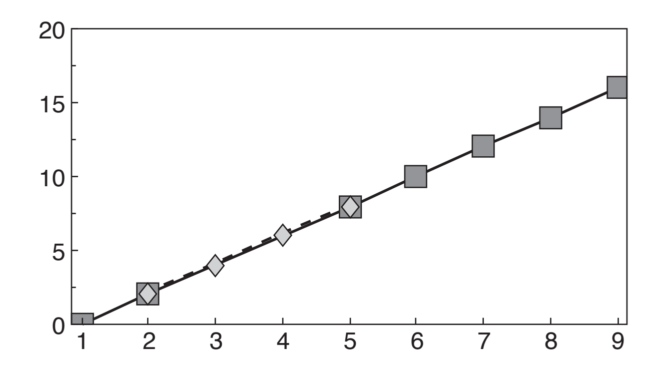
A related technique called *interpolation* attempts not to project forward, but to fill in the gaps in a set of historical data.

0, 2, ?, ?, ?, 10, 12, 14, 16

So, here's one guess at the missing numbers:

0, 2, 4, 6, 8, 10, 12, 14, 16

Again, it looks so easy when you describe the problem in a chart:



In practice, extrapolation and interpolation work well if the underlying formulas accurately describe the real forces that are shaping the events we observe. This approach works best for simple formulas—such as the compound interest on a certificate of deposit (provided that the depositor makes no unpredictable withdrawals!). Extrapolation and interpolation are also pretty reliable at describing the effects of natural laws, such as the flight of a ball through the air. Such predictable processes depend on just a few basic factors—the force and direction of the pitcher's throw, the force of gravity, the mass of the ball, and wind resistance. But when you leave the ballfield and venture into a complex field such as economics, the forces are many and largely unpredictable—at least by simple formulas.

The Lessons of Oversimplification

However complex or unpredictable the real forces may be, a line chart that describes the results can retain the *appearance* of simplicity, along with the false assurance that today's patterns will be somehow repeated in the future. It's a universal human tendency to try to see patterns,

which in most aspects of daily life is a mark of intelligence. But as you are perhaps beginning to suspect, this tendency can get you into serious trouble, chartwise.

Let's take a closer look at some specific types of trends you might see in chart data. I'll describe what these trends can mean, as well as how they can be used to distort rather than to accurately describe reality.

Beware of Your Average Chart

There are no pop quizzes in this book, but some of the stuff from Chapter 1 should come in handy right now. Recall from the example of the fibbing golfers that an arithmetic average, or mean, is the sum of the scores of previous games, divided by the number of games. Just to make things interesting, let's assume that the same golfer shot an unusually high 92 last time out:

$$\frac{79 + 81 + 78 + 80 + 76 + 92 \text{ STROKES}}{6 \text{ GAMES}} = 81 \text{ AVERAGE STROKES PER GAME}$$

In a line chart, the average is one type of trend, as shown in Figure 8.1. The individual scores form the peaks and valleys of a line plot, and the

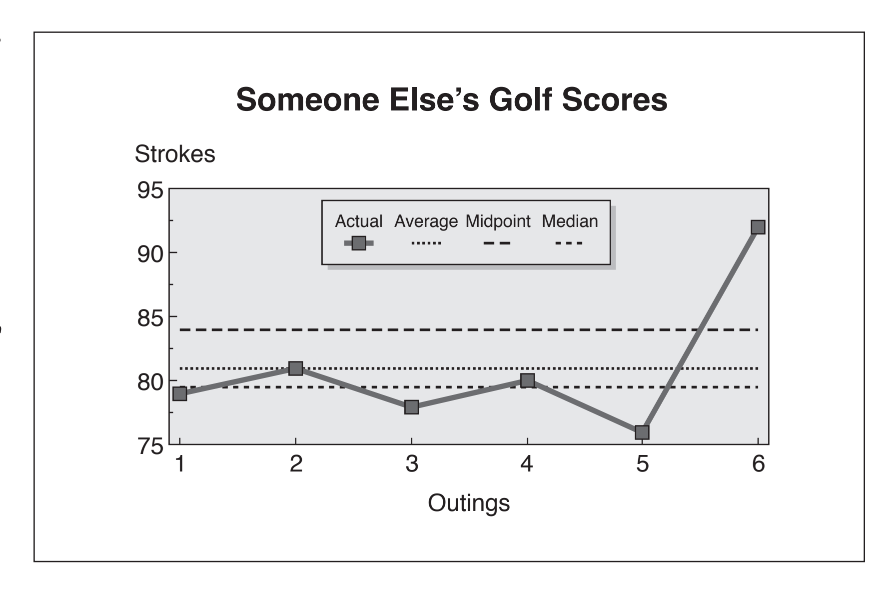


Figure 8.1 *The golfer's average score is shown here as a horizontal line that neatly cuts through the peaks and valleys of high and low scores. Also shown are the midpoint and the median—very different calculations that produce similar-looking trend lines.*

average is shown as a straight horizontal line that cuts neatly through them, leaving a symmetrical pattern of bumps on either side.

He Shoots a Middling Game

To unsuspecting audience members, the golfer's *midpoint* score might seem much the same as the average. To get the midpoint, you would take the average of only two values, representing a high and a low, by adding them and dividing by 2.

In the case our favorite golfer, the high was that pathetic 92 and the low was 76. The midpoint is

$$\frac{92 + 76}{2} = 84$$

For the charted result, refer back to Figure 8.1. Remember that the midpoint can be a better trend indicator if the values are wide-ranging, as they are here. For this golfer, the 84 midpoint might be a better predictor of his next game than his average of 81, which might be a better forecast of performance over a longer period of time.



Don't confuse *midpoint* with that other middling term, *median.*

Know What They Median

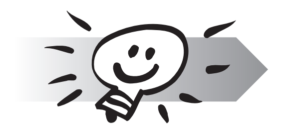
In a group of data values, the *median* is not an average, but one of the values that seems "normal." The median value in a set is located exactly in the middle—not in the middle of the range of values but in the middle of the number of items. So, there are the same number of values above the median as are below it. If you had a set of five scores, there will be two scores above the median and two scores below it. If you had an even number of data values, you would take the midpoint of the *middle two* as the median. For example, in a set of four values, you would take the midpoint of the second and the third values, so that there are two values above the midpoint value, and two below it.

Here again is the data for our golfer:

79, 81, 78, 80, 76, 92

There are six scores-an even number of items. The middle two scores are 79 and 80. (There are three scores below 80 and there are three above 79.) The median is

$$\frac{79 + 80}{2} = 79.5$$



Look back at the chart in Figure 8.1 to see how this method of calculating a typical score might distort the truth. *Using the median has the effect of ignoring exceptional highs and lows.*

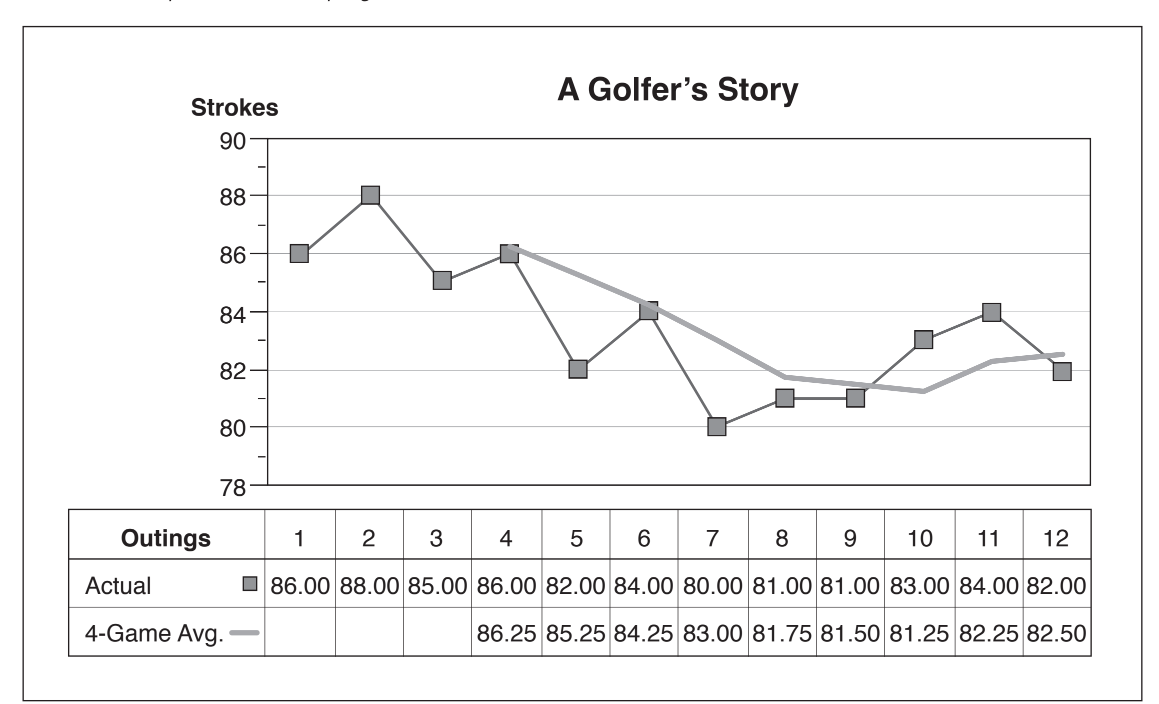
When in Doubt, Keep Moving

Recall that one of the golfer's last resorts might be a *moving average*, which would always include only the few most recent set of games-say, the last four. As each new game is included in the average, the oldest game drops off so that there are always just four games included. A chart of long-term performance, then, would show not only the actual scores, but also a series of four-game averages (see Figure 8.2).



Liars can use *average, midpoint,* and *median* interchangeably in chart descriptions, picking deliberately among these techniques to show the one that gives the most favorable result. A similar but perhaps subtler trick is to show a *moving average*, which might include only recent, favorable data, instead of using a longer-term *average*—and neglect to point out the "moving" part!

Figure 8.2 *The smoother line is the player's four-game moving average, which some people might regard as a better description of overall progress.*



Outings	1	2	3	4	5	6	7	8	9	10	11	12
Actual	86.00	88.00	85.00	86.00	82.00	84.00	80.00	81.00	81.00	83.00	84.00	82.00
4-Game Avg.				86.25	85.25	84.25	83.00	81.75	81.50	81.25	82.25	82.50

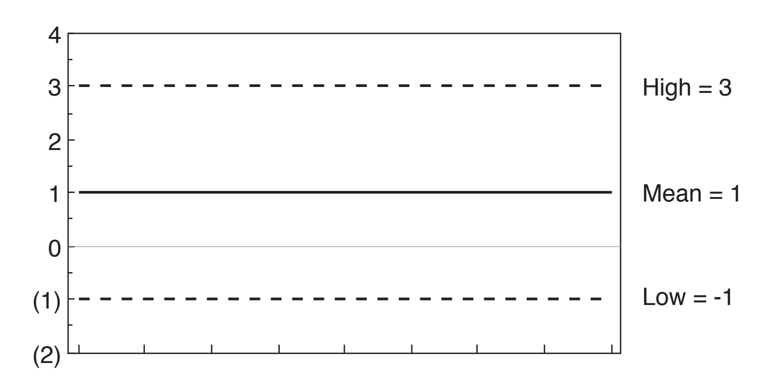
When Is Deviation the Right Road?

Statisticians are naughty mathematicians who, when kept after school, decided that they *liked* copying the same sentence a thousand times. Seriously, though, some of my best friends are statistics freaks—including many of you who were wise enough to buy this book!

More to the point, statisticians are world-class trend spotters. However, like any group of professionals, they have developed their own jargon, which can be downright intimidating to the average chart reader.

When speaking of the *mean* (arithmetic average), statisticians make the following distinctions, which may or may not have anything to do with what you want to know about a chart:

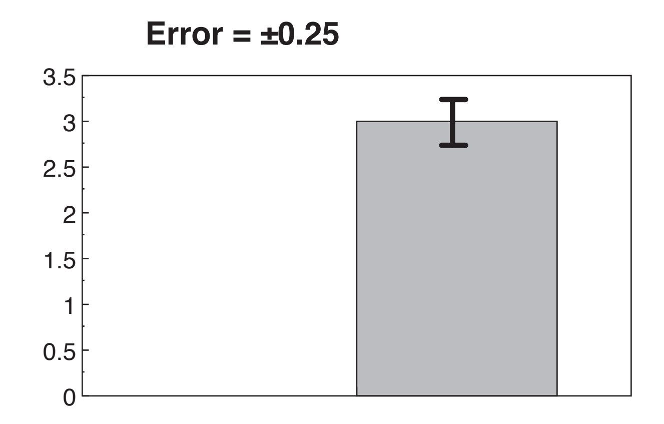
Standard deviation is the degree to which values in a series of numbers deviate from the mean. The standard deviation is usually shown as plus or minus the difference between the high and the mean, and the low and the mean. For example, if the mean is 1, the high 3, and the low –1, the standard deviation is ±2.



Population standard deviation is just a way of saying that all items in the data series were taken into account in calculating the standard deviation. The population standard deviation is used when the data series, or list of numbers, is small enough to be examined in its entirety.

Sample standard deviation indicates that only a typical selection, or sample, of data items was used to calculate the standard deviation. The sample standard deviation is used when the data series is very large, including perhaps millions of items.

In a bar chart, the standard deviation can be shown as *error bars*, indicating a typical range of values. More often, the presence of error bars indicates a degree of uncertainty: The chart maker is telling you that the actual value might lie anywhere within the upper and lower bounds of the error bar.





Liars are not usually honest enough to even think of using error bars, much less make the distinction between *population* and *sample* standard deviation. Scientists, who are supposed to tell it like it is, often show experimental results with error bars. The deviation might indicate, for example, limits on the accuracy of measuring devices.

How Can Regression Advance Your Career?

In the context of charting, *regression* doesn't have its psychoanalytical meaning—to return to a more primitive behavior—unless, of course, a chart maker is monkeying with the truth!

Regression is another trick of the statistician's trade for finding patterns in data. By trying different regression formulas, the statistician can fit smooth lines neatly through seemingly erratic sets of data points. The assumption is the better the fit, the more likely that the formula actually describes the process that is generating the data.

The process of applying this technique is called *regression analysis.* For math mavens, the algebra is in the sidebar, "Regression: The Gory Details." For the rest of us, it might be sufficient to know that regression calculations are built into business graphics and spreadsheet software.

Regression: The Gory Details

To generate a regression line, the data set must contain at least three data points. In practice, there should be many more to assure that any curve fit is meaningful. Different types of regression, defined by different equations, can be applied to the data to see which of the resulting lines produces the best fit through the points. The best fit is assumed to be the best model for predicting future results. Common regression types are linear, exponential, logarithmic, and power. Here are the formulas for each type:

Linear $y = a + bx$

Exponential $y = ae^{bx}$ where e is the base (~2.71828...) of the natural logarithm

Logarithmic $y = a + l_n(x)b$ where $l_n(x)$ is the natural logarithm of x

Power $y = ax^b$ where $a > 0$

In these equations, the dependent variable is *y* (the golf score in the example). There are three independent variables: *x, a,* and *b* (age, height, and weight).

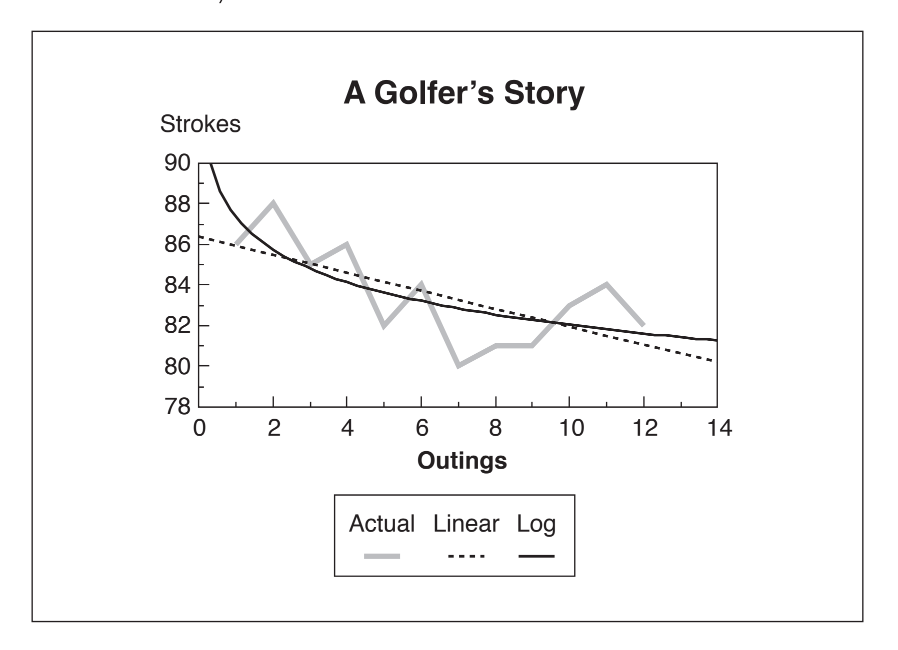
Figure 8.3 shows how two types of regression curves—linear and logarithmic—can be fitted to the same set of data points.

$$$$INVESTOR'S TIP$$$$

Regression analysis is more suitable for math geeks than general audiences, which might readily grasp a moving average but not these more exotic curves.

You can use regression analysis to analyze how a *single dependent variable* is affected by the values of one or more *independent variables*—for example, how a golfer's score (the dependent variable) might be affected by such factors as the person's age, height, and weight (independent variables). The formula that defines the link between the dependent and independent variables is called a *correlation*. In practice, given a large enough sample size, you won't find any meaningful correlation between golf scores and any of these factors. But that wouldn't prevent an overzealous statistician—or a dedicated liar—from using this math to justify a wrong conclusion!

Figure 8.3 *Two types of regression analysis have been done on the same set of data points: linear and logarithmic. The equation that produces the best fit is assumed to be the best model of the data. The linear regression emphasizes the downward trend. The logarithmic implies that the downward trend will level off eventually.*



Again, there may be no valid correlation between a golfer's scores and his or her age, height, and weight. But if you get a regression line to fit the data points, you will be strongly tempted to think that a correlation exists!

Regression analysis can be tempting for sincere forecasters and liars alike because it produces smooth line plots and curves that can appear to fit the data neatly. To the extent that there is a good fit, the underlying regression formulas would seem to describe the real forces that are shaping the data. The pitfall, as with looking for other kinds of trends in charts, is the potential to oversimplify, especially if the sample size is small or if some data points have to be excluded from the plot just because they don't fit.



The liar's trick with regression analysis is simply to exclude data points that don't fit some impressive curve. In the case of experimental data, for example, the excluded points are judged *anomalous*—statistics jargon for "I don't know what to make of these." Excluding anomalous data can be legitimate, though, as when making a large number of observations there will be some errors from inaccurate measurements or mistakes in transcribing the data.

A major hazard of computer-generated trend lines—as produced by regression analysis—is the temptation to rely on them simply because an electronic brain did the calculating. Before you jump to such a conclusion, remember—computers can lie just as readily as people can. Maybe even more so, since computers are literally without shame. Ethics and moral judgment aren't embedded in their programming. The subject of the next chapter is developing some healthy skepticism about chart data that is generated by computer.

9

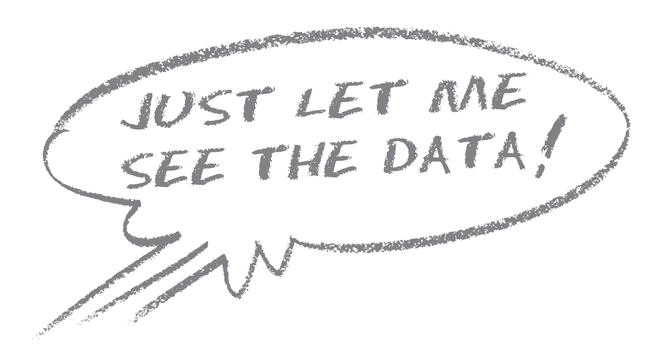
Tables and Spreadsheets

Or, The Socialite's Guide to Setting a Nice Table

As **every dinner-party impresario knows,** much of your success can depend on how you set a table. How you set (up) a table in a report or presentation can determine whether it will inform or confuse your audience. Furthermore, if you use a table as a worksheet for chart data, you will find that the arrangement of its columns and rows strongly affects your ability to generate meaningful graphs from the data.



The stuff in those graphic tables is particularly irresistible to avid number crunchers. To these information gluttons, a pretty sales curve is always suspicious; instead, they demand,



Their preference for seeing data laid out in tables is a cultural bias just as powerful as those notions of up, down, left, and right (which I expose shamelessly in Chapter 3). This bias comes not from geographic or ethnic background, but from the time-honored traditions of the accounting profession. I'll explore that notion in some depth in this chapter.

There's No Pattern on This Sheet!

Tables are not only boring, but also obscure. It's difficult to see patterns in the data when you are faced with a dizzying array of numbers. One of the reasons sales charts are curvy is for clarity!



Liars might prefer tables instead of charts if they are trying to obscure patterns in the data. If the array of numbers in a table is particularly dense and difficult to read, the presenter either is trying to hide the truth or is "graphically impaired" (more to be pitied than scorned).

As a general rule, you should do everything you can to use a graph instead of a table to present numeric results. Tables are better as worksheets than as formats for presentation. But, if your guests demand priority seating, I offer here some information etiquette for setting a nice table.

Tables and Sheets—You've Got It Covered

As I've said, notions about tables are rooted in financial accounting. Perhaps it's the close relationship between tables and spreadsheets that makes their arrangement by columns and rows seem so comfortable to some people. For accounting types, the spreadsheet—a kind of ledger—is the basic format for their working documents, or *workpapers.* (In the jargon of the trade, it's just a *sheet.*)

Spreadsheet Style—from Accounting 101

As a quick guide for the rest of us, the basic layout of an accounting-style spreadsheet is shown in Figure 9.1. The first row holds column headings. The first column holds row headings. The last column holds totals of each row. The last row holds totals of each column. The rest of the positions in the table hold numeric data. At the intersection of the last column and the last row is the *grand total*, which can be produced by summing either set of totals. Strictly speaking, all of the other totals

Figure 9.1 *In a conventional accounting spreadsheet, being able to generate the grand total by cross-footing—summing either down the last column or across the last row—provides a double-check on its accuracy.*

Third Quarter Office Expenses

	July	**August**	**September**	**Totals**
Wages	5,600	5,600	5,600	16,800
Supplies	436	327	408	1,171
Rent	1,200	1,200	1,200	3,600
Utilities	659	632	749	2,040
Totals	$ 7,895	$ 7,759	$ 7,957	**$ 23,611**

in the sheet are *subtotals* of the grand total. By accounting convention, subtotals should be highlighted by a single underline, the grand total by a double underline.

The accounting term for summing a series of numbers is *footing.* The ability to generate the grand total in two different directions is called *cross-footing.* Particularly when spreadsheets were done manually (that is, with sums done mentally, on an abacus, or with a calculator), cross-footing provided a quick check on the accuracy of the grand total: If you get different answers, at least one of the totals in the spreadsheet must be incorrect. If you get the same answer each way, the grand total is *very probably* correct. (You could have two or more errors, one compensating for another, but that's unlikely.)

Nowadays, a spreadsheet is an electronic worksheet generated by computer. Because the computer does all the calculating, there isn't as much need to prove the accuracy of the results by cross-footing.



Cross-footing is only one basic type of verification on the accuracy of a spreadsheet. At the conclusion of this chapter, I'll give you my Eight Great Steps to assuring the reliability of spreadsheet results.

Even though cross-footing is becoming rare, financially savvy people still expect to see totals in one of two places—at the right end of rows or at the bottom of columns.

Since a computer-generated sheet is much more fluid—and therefore changeable—than ink on paper, it can be much more convenient to show totals in the *top* row, as shown in Figure 9.2. In this arrangement, a kind of "perpetual ledger" can be created. In the example, the sheet is a check register. As each new entry for a check or a deposit is made in a new row at the bottom, the sheet grows longer. But the totals don't have to move. The updated numbers just overwrite the old ones in the top row. In a manual sheet, this would be messy and impractical. You'd have to use an eraser on the totals each time a new entry was made, quickly wearing holes in the paper. On paper, when you're at the bottom of a sheet, it's more natural to simply start a new one, carrying the balance forward.

Figure 9.2 *The electronic age makes alternate sheet layouts possible, but people have the entrenched notion that totals belong at the right side or at the bottom of a table.*

Totals		$244.92	$250.00	$1,005.08
Date	Reference	Checks	Deposits	Balance
			Start Bal. >	$1,000.00
01/03/95	721	$15.95		$984.05
01/05/95	722	$64.35		$919.70
01/06/95	723	$21.50	$250.00	$1,148.20
01/06/95	724	$40.00		$1,108.20
01/06/95	725	$103.12		$1,005.08



Presenting a table that has an unconventional layout, such as showing the totals at the top, can be disorienting to an audience. As you know from seeing liar's tricks with other chart formats, anything that runs counter to audience expectations can become another trick of the liar's trade.

Mixed Data Types—Another Cute Trick

$$$$INVESTOR'S TIP$$$$

You can use the Pivot Table capabilities of Excel to "put it right" by rotating the table to transpose rows and columns.

If a liar rearranges the totals in a table to confuse you, the trick will be fairly obvious. Frankly, the result is so odd-looking and unusual that most people wouldn't even think of trying it—much less tricking you with it.

But another trick of table layout is more subtle—and much more common. It has to do with mixing types or categories of data within the same column of a sheet or table.

Tables that are the easiest to read and understand follow the same rules used in constructing both traditional accounting spreadsheets and their modern-day variant—the computer *database* table. There's an example of a database in Figure 9.3.

Figure 9.3 *Because it is constructed according to strict rules that prohibit the mixing of data types within columns, a computer database table is very easy to read and understand.*

Monthly Professional Billings

Name	Job Skill	Area	Phone	Hourly Rate	Time	Gross Billing
Arnold, B.	Tax	818	555-9284	$200		$0
Delon, N.	Audit	818	555-2176	$200		$0
Enberg, A.	Audit	818	555-3876	$100	2	$200
Garcia, M.	Accounts	415	555-2019	$100	3.5	$350
Hofstra, G.	Litigation	213	555-3917	$200		$0
Kenworth, W.	Consult	310	555-1213	$100		$0
McMasters, O.	Tax	213	555-7856	$150		$0
Nanterre, B.	Accounts	213	555-2398	$100		$0
Ono, W.	Audit	213	555-1111	$200		$0
Otto, Y.	Audit	818	555-3331	$200	0.5	$100
Pulaski, E.	Accounts	213	555-6549	$150	1.2	$180
Tan, A.	Litigation	310	555-0432	$350		$0
Tso, J.	Tax	310	555-2857	$100		$0

Each entry in a database table is called a *field*. Each column heading in the first row is a *field name*. The field name describes one and only one type of entry, which applies to all the fields in that column. All the fields that describe one *entity*, such as a person, comprise one row. A row that describes an entity is called a *record*.

Look back at Figure 9.1, the traditional accountant's spreadsheet. Except for that last row of totals—which often holds numbers that are used only for purposes of cross-footing—the sheet follows the rules exactly for construction of a database table.

So, both the accountant's spreadsheet and the computer database table show data according to rules that fit audience expectations, including the requirement to keep different categories of data in separate columns.



Intermixing data categories in the same column of a table can be another disorienting trick of the accomplished liar.

A table that violates these rules is shown in Figure 9.4. You can see right away that it's a kind of apples-and-oranges layout. This graphic is actually not a table but a *list* of dissimilar (but perhaps related) items, presented in tabular form. A good table is useful in part because it invites *comparisons* among similar entities—usually, one entity in each row. But in the liar's example, no meaningful comparison is possible among the rows. The trick is precisely to *prevent* the audience from making comparisons.

In this example, comparisons are difficult within a data category (such as Test Scores) because each day's score would have to be shown on a different slide or page of the presentation. Notice how much easier it is to compare the daily scores when they are presented in the same table (see Figure 9.5).

Figure 9.4 *This table violates the rule about showing only one data type in each column.*

LIAR'S TRICK



Day	Monday
No. of Students	32
Room Temperature	75.6
Test Score (Avg.)	82.0

Figure 9.5 *This table keeps data categories consistent within each column, making it possible to make comparisons among rows.*

Student Testing Environment

Day	Students	Temp.	Score
Monday	32	75.6	82.0
Tuesday	28	78	81.5
Wenesday	29	73.4	80.9

Information Etiquette for Table-Setters

Given the rules for constructing tables that people can readily understand, here are some tips on using tables (and electronic spreadsheets, if you must) in presentations.

Use tables to make your explanations more concise.

In a report narrative or a speech, you can avoid lengthy explanations of key relationships by summarizing them in a table. You can then use fewer words to simply highlight specific entries in the table or to draw overall conclusions about relationships among the entities. Be guided by the fact that businesspeople are always in a hurry—they won't be reading your report or attending your presentation for pleasure.

Don't present data in a table or spreadsheet if a graph can show the same idea.

Raw data is seldom as interesting as a chart, which can make a trend or relationship understandable at a glance. If the source data must be shown, consider presenting the numbers in a table beneath the graph:

	Aardvark	Aero	Acme	Able	Archer	Argo	Ace
Last Year ■	93	71	99	94	121	105	60
This Year □	124	83	162	95	143	231	78



Notice that the *data table* shown above contains not only the graph data but also a *legend,* or data series key, with color codes for each row. Legends are discussed in Chapter 12.

When covering dissimilar things, summarize them as topics in brief lists, not in tables.

Use bullets when presenting topics and key points, especially if you are listing topics that will follow as subtitled sections of your material. Use a numbered list instead if you a presenting a sequence of steps. Professional presenters call this technique *signposting,* because it can guide the audience through the outline structure of your presentation.



Limit the number of columns in a table.

The controlling factors on the number of columns in a table are page or screen width, margins, and text size. So, you can usually increase the size and therefore the legibility of the text if you use fewer columns. You might productively eliminate a column used for explanations by placing the information instead in a footnoted list at the bottom of the table.

In a printed report, try to keep a table on a single page, close to its explanatory text.

A page break in the middle of a table makes it that much more difficult for your readers to follow, especially if they will be making comparisons among the rows. Instead of breaking a table, consider placing it on a separate page, even if it must be located farther away from its explanatory text. Give the table an exhibit or figure number, and include that reference in the text.

$$$$INVESTOR'S TIP$$$$

See Chapter 13 for more tips on building and checking the consistency and accuracy of tables in financial reports.

Justifiable Justification—Here Are the Rules!

The rules for justifying, or aligning, table entries are pretty standard. If you follow them, your tables will not only be easy to read, they'll also look neat, which always gets you extra points in my book. The rules are:

- Align alphabetic text, such as labels, left:

Acme	
Arrow	

- Align numeric values right, usually on the rightmost digit:

Acme	253
Arrow	4,841

- If numeric values are decimals, align them on the decimal points:

Acme	253.07
Arrow	4,841.372



You needn't be concerned about aligning digits on their decimal points if you use the same number of decimal places for each entry in a column and align them all on the right.

- Optionally, center the headings in their columns:

Company	Index
Acme	253.07
Arrow	4,841.372

Publishing Spreadsheets, If You Must

All the tips about tables also apply to computer-generated spreadsheets—which, for the most part, are just big tables. You should try to avoid including spreadsheets in your reports because they are, well, just plain boring. Furthermore, your purpose should be to summarize information, and spreadsheets typically have too much detail. Again, you might communicate your message better in a graph that uses just a few critical rows from your spreadsheet as data series.

But, if you must, you must. Sometimes they insist on seeing the numbers.

The rest of this chapter is not so much about laying out the numbers visible on sheets but about the calculations underlying those numbers. It's crucial to the truth of your tables—as well as to the veracity of the charts you make from them—that you get a passing grade in spreadsheet logic. If you don't understand how the numbers are being generated, you leave yourself open to all kinds of unintentional errors—and to intentional distortions made by liars who aced the course!



The spreadsheet examples in this chapter refer to Microsoft Excel, but the concepts apply equally well to OpenOffice Calc and other spreadsheet programs.

Recall from Chapter 1 that a liar's first step in creating deceptive charts can be to manipulate the data from which the charts are generated. Well, as Big Mama says in *Cat on a Hot Tin Roof*, the biggest problems in life are usually under the sheets!

Finding Errors in Spreadsheets

Computers are speedy and meticulous calculators, but their programs know nothing about the larger business problems they are solving. They therefore cannot catch many kinds of human error, particularly errors in the way problems are described in math formulas.

Spreadsheets: The Crash Course

You needn't skip the rest of the stuff in this chapter just because you haven't yet experienced the joys of electronic worksheets. Although rocket scientists rely on spreadsheet technology, you don't have to be one to use it.

The columns of a sheet are marked with letters, starting with **A**, across the top:

A	B	C	D	E	F	G

Rows are identified by number, starting with **1** and running down the left side of the sheet:

1	
2	
3	
4	
5	
6	

The intersection of each column and row forms a small box. Each box is a storage location for once piece of stuff, such as a number or text. Each box is called a *cell*, and each cell has an *address*. The address of a cell is the letter of its column and the number of its row, with the column always coming first. So the address of the top left cell in the sheet is **A1**. (I don't know about you, but I find the term *cell* a bit chilly. If it makes you more comfortable, call it a "cubbyhole.")

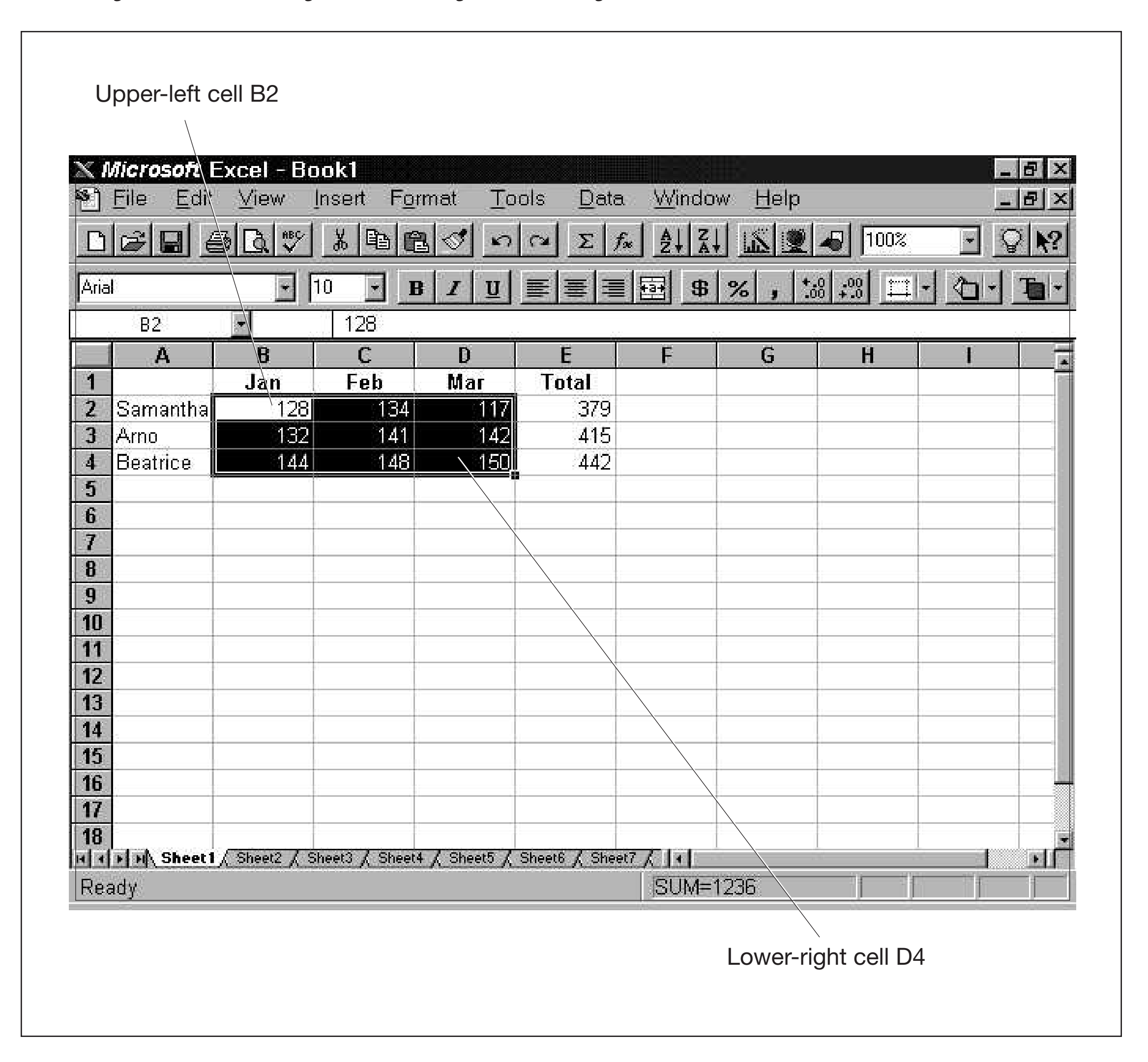
	A
1	

Each cell has its own address, and no two cells have the same one. If you tell the program that you want the stuff you stored in cell D5, it will quickly fetch the contents of the box at the intersection of column D and row 5.

Picking Your Favorite Cell Block

It's often necessary to refer to a block of cells, such as the numbers in a data series for a chart. In spreadsheet jargon, a *range* is a rectangular block of cells defined by its first cell and its last cell. The first cell is in the upper-left corner of the range and the last cell is in the lower-right corner of the range. Figure 9.6 shows the range **B2:D4**. When you specify a range, you are selecting the contents of its cells. The colon (**:**) means "and all cells through and including." (Some spreadsheet programs use a different notation: **B2..D4**.)

Figure 9.6 *Here's the block of cells (or spreadsheet range) B2:D4, defined by its top left corner (cell B2), and containing all the cells through and including its bottom right corner (cell D4).*



Name Your Range!

There's a much more understandable way of referring to ranges—you can give them names. Of course, this is especially handy if the name you pick describes the data that's stored there, such as **SALES**.

(Continued)

Spreadsheets: The Crash Course (Continued)

Letting the Computer Do Your Math

Instead of holding a piece of data, such as a number or some text (called a *label*), a cell can hold a little math problem, or *formula.* A formula is a calculation that the computer will do, saving you the effort. Here are the elements of a typical formula:

- **Equal sign** A formula usually begins with an equal sign (**=**). It tells the program, "Do this math!" (In some programs, a formula must begin instead with **@** or **+**.)
- **Function name** A function name—**SUM**, for example—is shorthand for a more complicated formula that the program already knows how to do. **=SUM** (or **@ SUM**) adds a series of numbers; **=AVERAGE** (or **@AVG**) averages them.
- **Argument** Range addresses or range names enclosed in parentheses form an *argument*. With functions such as **SUM** and **AVERAGE**, the arguments describe a block of cells that holds numbers. An argument tells the program where to find numbers to give to a function.

HERE'S A FORMULA:
=SUM(B2:B7)

This formula simply means, "Sum the numbers in the range from cell **B2** through and including **B7**." The program does the work and puts the answer in the cell that contains the formula. The use of this formula in a sheet is shown in Figure 9.7.

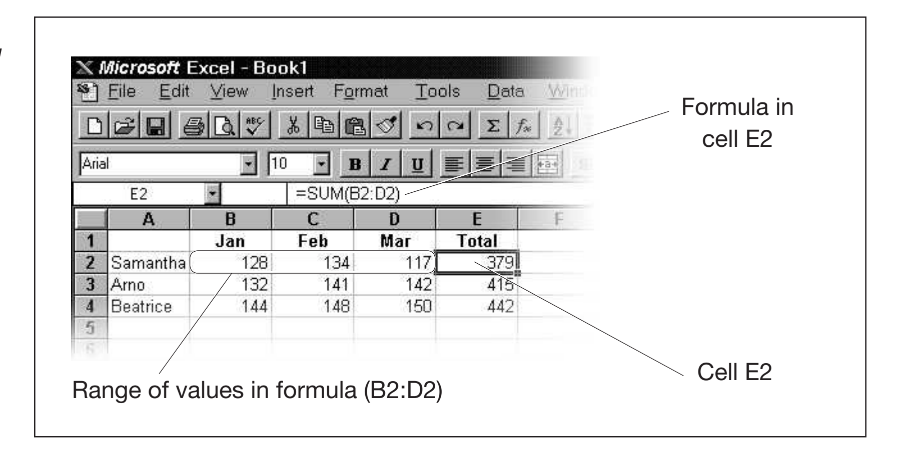


Figure 9.7 *Here's how a formula that uses SUM might be used to add a row of numbers in a spreadsheet.*

That's it! That's all even rocket scientists need to know about how spreadsheets work.

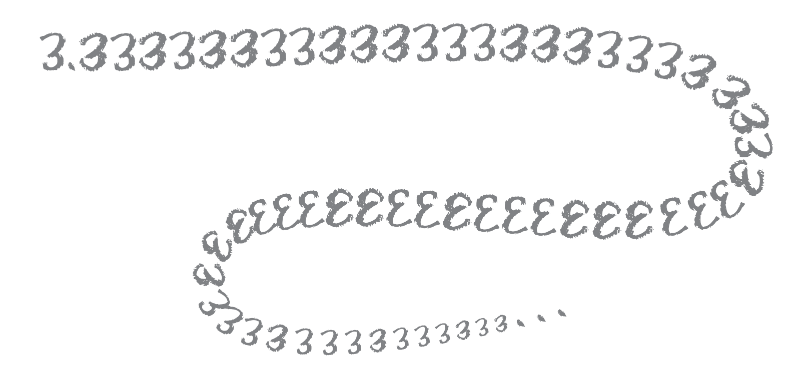
Spreadsheet errors can be intentional—deliberate efforts by dedicated liars—or unintentional—mistakes by spreadsheet makers whose logic is flawed. *By far the majority of errors are not committed on purpose.* I don't mean to say that there aren't many liars, there are just far more dumb mistakes! And, in a way, that's a scarier situation, because a lot of people out there in computerland are generating misleading data with the best of intentions.

For the rest of this chapter, I'll explore some common sources of spreadsheet error which can affect both tables and charts. I'll describe two basic mistakes: 1) picking the wrong formula, and 2) using flawed logic. There are many variations on these, but you'll get the idea. Then, I'll conclude by giving you my Eight Great Steps you can take to check a sheet for all types of errors.

Working with Decimal Values

Formulas that use decimal values can produce errors that are difficult to spot, even for experienced users of spreadsheet software. Getting a reliable answer involves picking the right function as well as understanding its effect on the rest of the math in a spreadsheet.

Decimal values are problematic because a spreadsheet cell can only hold a specific number of digits. This fact, in itself, is a potential source of error, since some decimals could go on forever. For example, 10 divided by 3 equals



That ellipsis (...) means "and so on," so what you see is just an approximation, despite all those digits.

Most spreadsheet programs provide a variety of readymade functions you can use to manage decimal values that have a lot of digits. These functions all work differently, so picking the right one is essential to getting correct results. Two particularly troublesome functions—*rounding* and *truncation*—give similar results that can create big differences.

Rounding

The function ROUND does the seemingly simple job of rounding off a long decimal value to a shorter, more manageable one. In so doing, it can adjust the last digit of the result so that it better approximates the original, longer number.

A rounding formula typically takes two arguments—the decimal value you want rounded and the number of decimal places you want in the result:

```
=ROUND(B2,2)
```

In this example, if cell B2 held the value 1654.9876, the result would be 1654.99 (with the last digit rounded up). ROUND can be the most accurate way of dealing with unruly decimals—provided that the result can also be a decimal.

The second argument you feed to ROUND can be zero or a negative number. If the second argument is zero, the program will round the result up or down to the nearest integer. For example, an argument of 0 would round 34.9 to 35. If the second argument is negative (less than zero), the program will round the number to the left of the decimal point at the position indicated by the number. So an argument of –1 would round 34.9 to 30.

Truncation

A function that works very much like ROUND is TRUNC, which truncates, or just chops off, some or all of the decimal digits. Like

ROUND, TRUNC takes two arguments—the number you want to truncate and the number of decimal places you want in the result:

=TRUNC(B2,1)

In the formula above, if B2 held the value 1654.9876, the result would be 1654.9. If you omit the second argument, it is assumed to be zero, in which case the program will throw away the entire decimal portion, leaving an integer (whole number with no decimal places) as the result.

Sources of Error from Approximations

Serious errors from ROUND and TRUNC stem from the fact that they produce *approximations* as results. The differences in a single answer might be small, maybe insignificant. The errors grow larger, though, when the results of these functions are then used by other functions, especially in multiplication. If a small error is multiplied many times, it can become big enough to matter.



In a famous case of "Truncation Abuse," a bank employee noticed that a computer program used truncation instead of rounding to calculate interest on savings account deposits. The employee modified the program to send the discarded part of the truncation to his own account, eventually accumulating a considerable sum. (Before you get any ideas, the Feds are wise to this one!)

Other Similar Functions

Spreadsheet programs have many, many built-in functions—too many to cover in this modest book. But ROUND and TRUNC are representative of lots of other groups of functions that do similar things but can produce significantly different results. For example, you might pick the function INTRATE to calculate the rate of return on an investment. INTRATE *might* be correct if the investment is a bank certificate of deposit. But if the investment is a bond, you should use YIELD, and if it is a U.S. Treasury bond, TBILLYIELD. These are functions which have formulas specifically designed for those types of investment.



Calculating investment return is particularly prone to errors (or intentional misstatement) due to repeated multiplication because the return will be *compounded,* or reinvested periodically. So, the result of even a small margin of error will be literally compounded, growing ever larger over the period of investment. If a liar were trying to sell you an investment, he or she might try to overstate the return, which can be done easily with such "funny" math.

Running in Circles

Another source of spreadsheet error that can be particularly troublesome is the *circular reference*. A circular reference is the result of a formula that refers to itself.

Direct Circular References Are Easy to Spot

One type of circular reference is *direct*. This happens when you enter the address of the current cell in the formula that it holds. For example, cell **A1** might hold this formula:

=A1+1

Unless it appears anywhere *but* in cell **A1**, this formula describes a logical impossibility. When you enter the formula there, the program replies, "Cannot resolve circular references." However, the program will go ahead and insert the value 0 (zero) in the cell. At this point, your sheet contains a logical error, which isn't obvious from that innocent-looking zero.

Indirect Circular References Can Be Hidden

The indirect kind of circular reference looks less like a mistake, and is therefore both easier to make and much more difficult to trace. For example, cell **A1** might hold the formula

=B1+1

This, in itself, is perfectly fine. No circular reference here. But here's what's in **B1**:

=A1 * 2

As you can see, **A1** has a formula that refers to **B1**, which in turn has a formula referring back to **A1.** Although the reference is more roundabout, it is circular nonetheless. Once again, the program replies, "Cannot resolve circular references," but goes ahead and inserts a zero in the cell.

Circular references are not necessarily wrong. For some kinds of problems, it can be productive to keep the computer running in circles, so to speak. The trick is to limit the number of circular calculation passes, or *iterations*. Two criteria can control when the circular calculations will stop: *maximum number of iterations* (say, 100 passes) and *maximum change*, whichever occurs first. The maximum change value becomes important if repeating the calculation causes the result to come closer and closer to some value that it never quite reaches. For example, calculation might stop when the next pass produces a difference smaller than 0.001.



Circular references are just one example of flawed spreadsheet logic resulting from one formula using the results of another. The inter-relationship of formulas in a sheet can get very complicated, and that's why makers of spreadsheets must not only be careful designing their sheets but also must devise tests to make sure their formulas work as intended.

Eight Great Steps to Spreadsheet Reliability

Here are some basic steps you can follow to ensure that your spreadsheet (or one designed by a suspicious third party) produces reliable results:

1. Pick the Right Function for the Job

Beware of differences between functions that have similar purposes: INTRATE for securities, YIELD for bonds; ROUND for rounding, TRUNC

for truncation; AVERAGE for arithmetic averages, and MEDIAN for the median value.

2. Watch for the Kinds of Errors the Computer Can't Catch

Avoid division by zero, or by any reference that might result in zero.

Beware of any rounded or truncated value that is used in other calculations, particularly multiplication. Inaccuracies will get worse each time an approximated value is multiplied.

Avoid circular references unless you need to solve a problem that requires many recalculation passes (iterations).

$$$$INVESTOR'S TIP$$$$

Auditing your worksheet formulas is a crucial quality-control step in Excel. Apply test data within and outside of expected ranges. Also be continually suspicious of the reliability of automated links to external data sources that are subject to change.

Be careful when doing calculations with dates, such as subtracting dates to find out how long an invoice has been unpaid. Computer systems and accounting departments vary in the ways they handle dates: for example, does the date 09/06/2049 mean September 6 (U.S. convention) or June 9 (European convention)? Do *serial date values*—the format in which your computer actually stores and works with dates—begin at the year 1900 (IBM PC convention) or 1904 (Apple convention)? Is the accounting year based on the normal 365-day calendar or on a special 360-day year with all 12 months having just 30 days each? Does the accounting year start in January or is it a fiscal year that starts in, say, July?

3. Make Sure to Give a Function Everything It Requires

If you don't enter the arguments as required, the program might be smart enough to reject it. For example, it may respond, "Too few arguments," or "Too many arguments." Or it might put a coded error message in the cell instead of a numeric result. But it has no way of knowing if the cell addresses you gave it actually hold the numbers you want to use.

Be sure to supply the correct *keywords* (special-purpose literal text strings) for the functions that require them, such as **INFO**. Enclose keywords in quotation marks (" ")and spell them correctly. For example, the formula

=INFO("MEMAVAIL")

will tell you how much free memory is in the computer.

Enclose text values used as arguments in quotation marks:

"JANUARY"

If you omit the quotation marks, the program will think that the item is a range name and will try to find its value.

In arguments, you can use parentheses to control the order in which calculations are done:

(2*4+3)=11

(2*(4+3))=14

Parentheses must be used in matching pairs. If you get lost among such *nested* pairs of parentheses, count them. If you have an odd number, you know you have not used the parentheses correctly. But if you have an even number, you can't assume that the rest of formula is correct. Always make sure to check sets of parentheses carefully.

TRUE and **FALSE** are logical values (equivalent to **1** and **0**), not text. Do not enclose them in quotation marks when you use them in formulas or when you type them as values in cells.

4. Verify All References to Data

Make sure that arguments refer to cells or ranges that hold the values needed by a formula. This is probably the most common source of errors from formulas—especially from formulas that *appear* to work because they produce values instead of error messages.

Use range names, such as

SALES EXPENSES PROFIT

for clarity, but don't assume that the name refers to the correct range until you recheck it. Beware of copying or moving cells or ranges within

a sheet—the move can change the references in any formulas it contains. (The program might make adjustments for this automatically, but the results won't necessarily be correct.)

5. Double-Check Your Math

When you're done building a spreadsheet, make a copy of it for testing purposes. (Testing might produce more errors in the sheet, so always test a *copy*.) Enter sample data and note the results. First, use stuff that produces answers you already know or can calculate manually. Then, use data at extremes (high and low anticipated values), and even try data that's out of bounds. Use data of incorrect types, such as a date entered as text where a date value is expected, as well as of correct types.

Watch for error messages and incorrect results. Revise your formulas until the sheet works as expected.

6. Ask: Do These Results *Look* Reasonable?

Generating graphs from ranges in a sheet can be a quick visual check on the probable accuracy of results. (The accounting term for an estimate of probable accuracy is *reasonableness*.) Besides, charts are prettier, and learning how to make them makes good reading.

7. If You Make Changes to the Sheet, Recheck and Retest It

Remember that a spreadsheet program will usually change the cell references in formulas if you copy or move ranges, or if you delete or add columns or rows. Inspect and recheck the formulas after you've rearranged the sheet to be sure that they are referring to the correct sources of data.

8. If You Think Your Computer Goofed, Think Again!

Spreadsheet software can catch the kinds of blatant errors you might make by entering a formula incorrectly. But if your formulas are correctly composed (have correct *syntax*), the computer will slavishly do all your math without complaining. No computer program, however

sophisticated, is smart enough to detect errors in your understanding of underlying business problems. Your spreadsheet might appear to work, but only you can judge whether it works right.

Charting Your Way to a Brighter Future

The old proverb that a picture is worth a thousand words goes for numbers, too. Don't publish data in a table or in a spreadsheet that can be shown more clearly and persuasively in a chart. Spreadsheet software has built-in graphing capability—so what's your excuse?

The most important consideration in making a chart from a spreadsheet has nothing to do with computer software: You must exercise some careful judgment in selecting which data will be charted. No graphics software yet invented can make sure that the input data actually describes a meaningful trend or a key relationship. That's why programs that are advertised to generate graphs with a single command or click of a button aren't necessarily a help. You might be able to make a chart quickly (and there's no reason why you shouldn't!), but your understanding of the data, not the ease of making pretty pictures with it, will determine whether the results will make sense to anyone.

A rule of thumb is to be downright picky about the data you use for charting. In any sheet, there is almost always too much information to make a single, understandable chart. Out of the whole sheet, you usually need to find just *one or two rows* of data that tell a story.

10

Technical Analysis Basics

It's Anything But Fundamental

THE **focus in this chapter** is on *technical analysis* of financial market data. Now, I don't include it because financial charts are inherently deceptive. No, my cautionary advice has to do with the human tendency to see patterns and derive meanings from them. With financial analysis, asking the "why" of market ups and downs is a pitfall. Technical analysis of charts is all about the geometry and the math of the patterns—not the causes of those market fluctuations.

Those causes are the subject of study for a different discipline—*fundamental analysis.* That approach deals with value judgments about how a company is performing and why. Such judgments follow from circumstantial questions that range from the general to the specific.

- General questions include: Is the need for the company's products rising or falling? Are the products priced right? How intense is the competition? Is the company capitalized sufficiently? How capable is its management?
- Some specific, more quantitative questions would be: Where is the current stock price in relation to the company's historic highs and lows? Has the company made its recent revenue forecasts? What is the ratio of the price of the company's stock to its earnings? How is the company's overall sector performing?

These questions arising from fundamental analysis represent a traditional approach to investment decisions. By contrast, technical

The Future of Technical Analysis Research

I'll qualify my assertion that technical analysis has nothing to do with causes. That's the state of the art. No one knows for sure what market forces and circumstances give rise to a particular chart pattern. Oh, yes, there are statistical clues and metric indicators. But that's all math, not sociology or geopolitics. However, I'm not saying that chart patterns exist independently of those causes. Some scholars are applying statistical analysis and data mining to find explanations for those patterns. But any explanations are highly theoretical. Patterns may possibly have to do with the mathematics of chaos theory or with the group psychology of "herd mentality." It may be possible someday to read financial charts the way a heart specialist reads an electrocardiogram. But that day is not yet here. The complexity of the challenge is more akin to explaining what specific climatic conditions caused the patterns of clouds you see in the sky over your town today.

---

analysis focuses on rules of thumb developed from observations about patterns and their tendency to recur. The causes behind those patterns are irrelevant. And, given the current state of the art, largely unknown.

In fact, looking for market causes in chart patterns can actually confuse the process of using technical analysis to predict what the market will do next.

So why does a discussion of technical analysis belong in this book? First of all, many people who study charts are interested in this topic, including me. I'm no expert on financial trading or charting, but this chapter summarizes what I've learned about it. And what's relevant here is that I've applied the principles in this book to my study of this topic. Remember, *How to Lie with Charts* is a how-*not*-to book. Focusing on certain aspects of technical analysis can help you develop a better critical eye for interpreting—or making—all types of charts.

And in the broader context of financial reporting, this critical eye can help you make better business decisions from analytical tools.

Chart Patterns and Interpretation

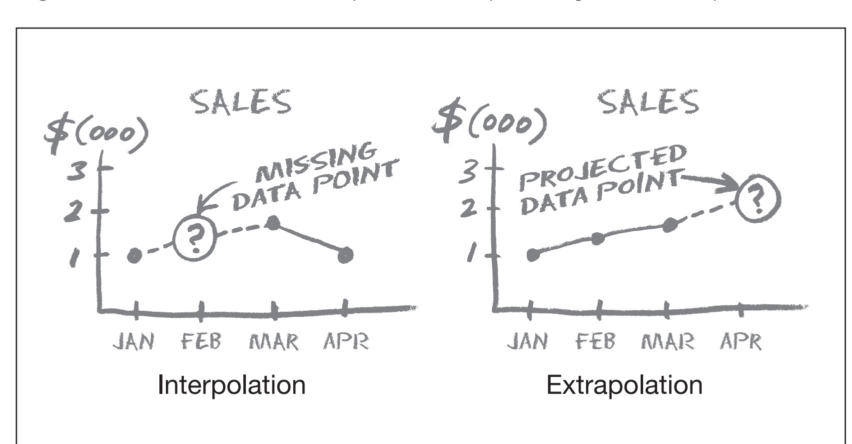
All the math fundamentals of *interpolation* and *extrapolation* amount to reading between the lines (Figure 10.1). It's another way of finding meaning in patterns, which the human brain does more or less automatically.

And that's because human beings are "meaning-making machines." When our ancestors ventured out into the jungle, we were looking for things that would eat us, things we could eat, and other people we might want to copulate with. And so we're constantly looking for patterns and making meaning out of them. And our success at recognizing patterns has been one of the reasons for human survival. It's a basic skill, and you can't turn it off.

So we are programmed to find patterns. We automatically find meaning in every pattern we see. But if you start reading meaning into stock market charts, you're really going down a slippery path.

Figure 10.2 is a graphical snapshot of the stock market. But this is not the market any more than a photo of you is the "real" you. This chart is a representation of the stock market—an abstraction, a data reduction. But it is not reality. No chart is reality. It's a means of interpretation. It's a

Figure 10.1 *As discussed in Chapter 8, "trendy thinking" can be risky.*



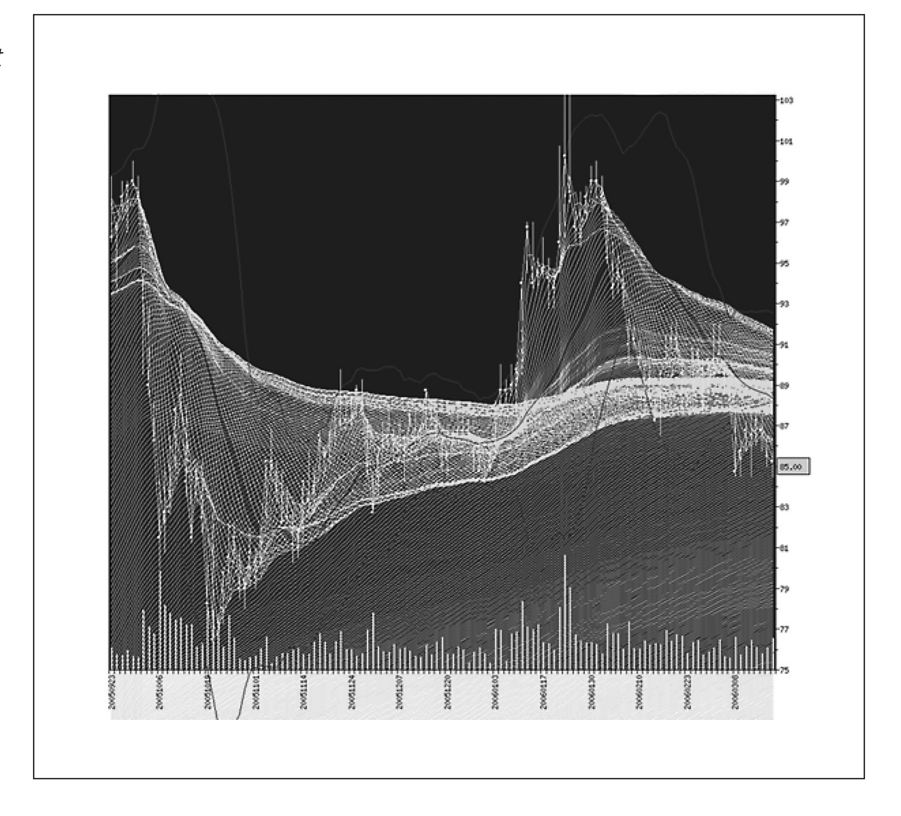


Figure 10.2 *A chart is no more the stock market than a photo of you is the "real" you.*

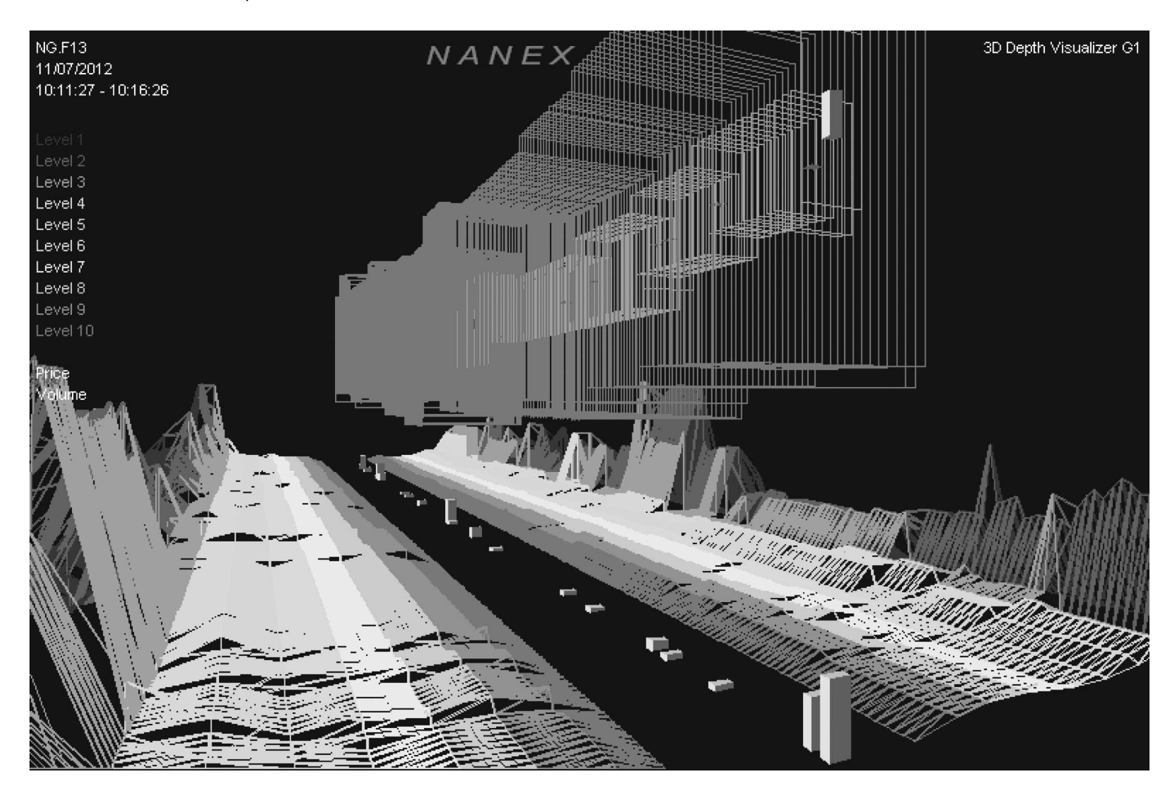
tool, a basis for presenting some information that is a partial and highly selective view of reality.

In this case, the picture is a chart of price over time. As you would expect, and according to convention, the dollar value of the stocks traded fluctuates up and down on the *y* axis, and time flows from left to right along the *x* axis.

Regarding technical analysis, this is a typical chart because it presents a series of months rather than hours or days. That's significant—because technical analysis generally does not deal with the types of short-term patterns you'd see within a trading day.

As to the short term, Figure 10.3 is a picture of a data feed showing how the market is performing from moment to moment during a single day—more or less in real time. Is this a better picture of reality? No, it's simply a different set of arbitrary assumptions. And in some cases it may

Figure 10.3 *A real-time data feed of stock market prices does not necessarily have more realism than a longer-term view, and it is generally not suitable for technical analysis, which focuses on durations of weeks and months. (Courtesy Nanex)*



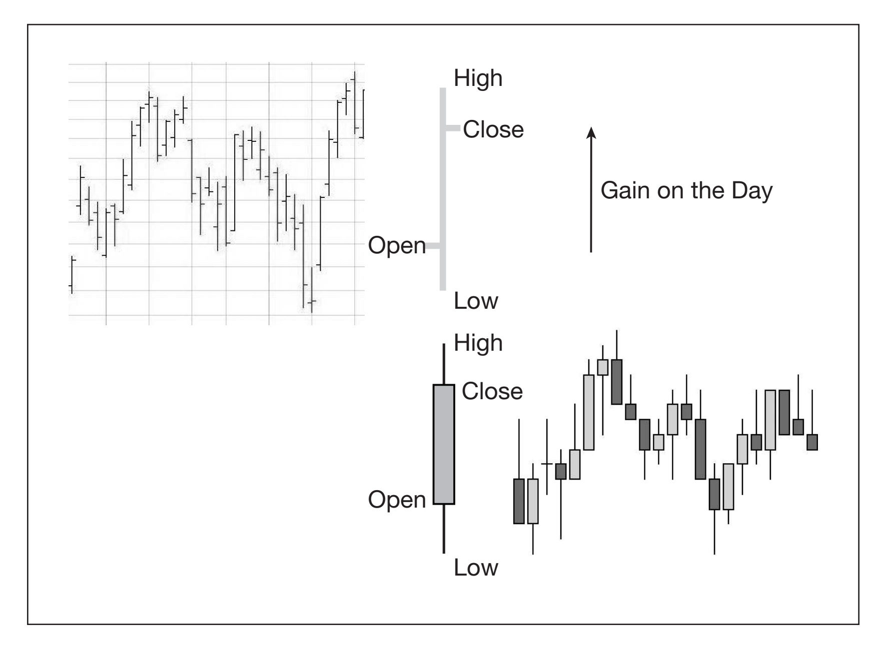
be *less* useful because it's not *enough* of a data summary, not enough of an abstraction. It's not giving us that bird's-eye view of how our favorite stock or index is performing over time. The real-time feed is more like raw data, and raw data is much less useful for interpretation.

Reading Candlestick Price Indicators

A mainstay of financial charting is the candlestick price indicator, which has four data points, as shown in Figure 10.4: High, Low, Open, and Close. Specific graphic styles vary, but the vertical bar always represents the span of prices in a single trading session, or day. If the bar is green, the price closed higher than it opened (Close > Open), resulting in a



Figure 10.4 *Here are two styles of candlestick price indicators. One candlestick represents the range of prices of a given stock (or index) on a single trading day. Technical analysis charts closing prices over longer spans.*



gain on the day. If the bar is red, the opposite is true and the price had a *loss on the day*.

Again, technical analysis generally does not focus on daily fluctuations. Instead, the chart under study shows the fluctuations in the daily closing price over a period of weeks or months. The chart may show the prices of a stock for a single company (or price of a particular commodity), or the chart may show how an index, or a "market basket" of stocks, is performing. Two commonly used indexes for the New York Stock Exchange (NYSE) are the Dow Jones Industrial Average (DJIA) and the Standard & Poor's 500 (S&P 500). Charting an index can tell you how well the market is performing in total, highlighting trends or cycles that might be affecting most, if not all, stocks.

Figure 10.5 charts trades of a particular stock over a period of several years. By convention in all types of financial charts, increases or decreases in dollar value and volume are plotted against the *y* axis. Time flows from left to right along the *x* axis.

Figure 10.5 *This line plot shows the variation in stock price. The vertical bars show the volume, or corresponding number of shares of the stock bought or sold.*



Drawing Support and Resistance Trend Lines

Figure 10.6 shows a simple approach to technical analysis—drawing straight lines on the chart to highlight trends. It's all about calculating slope, or value divided by time. These are simple *projections* of future price levels. The bottom lines connect the low points leading into and coming out of a trough. These are *support* lines, indicating the lower bounds of a trend. The line on top connecting the two peaks is a *resistance* line, indicating an upper limit on prices. With a chart pattern, the general rule is that the price will probably not fall below the support line nor rise above the resistance line. If it does and keeps going, it's called a *breakout* of the pattern, which may be the start of another pattern.

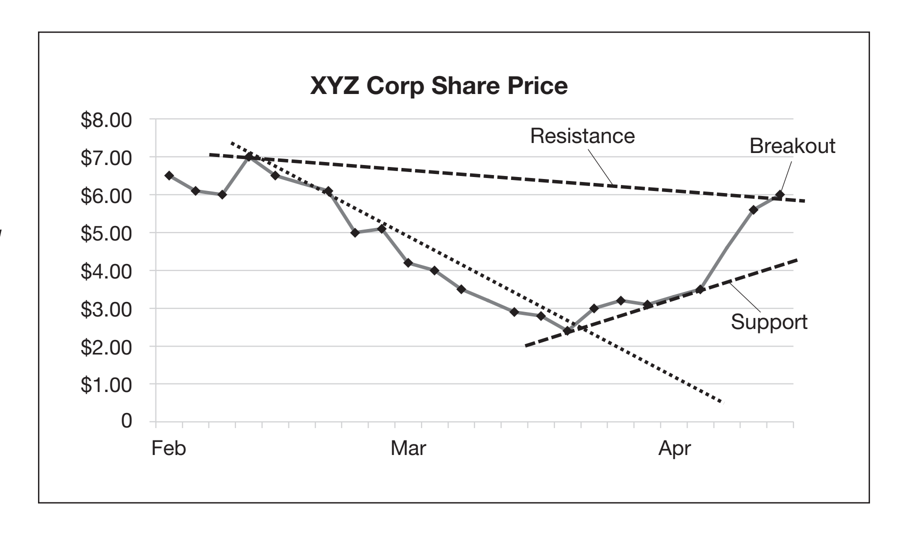


Figure 10.6 *Simple projections of trends find the slopes of lines drawn through highs (resistance) or lows (support). Deviation from a pattern, diverging from a trend, is called a* breakout.

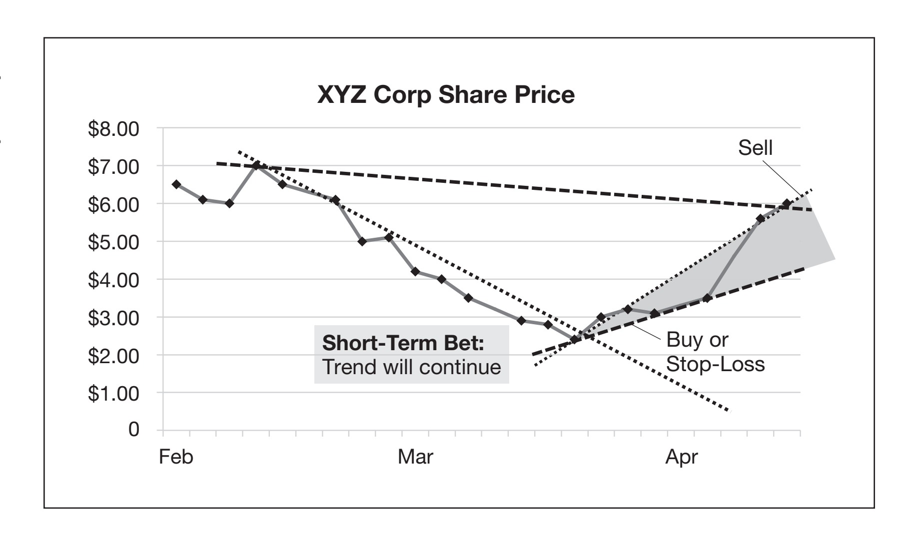


Figure 10.7 *In this example, the trend lines on the right mark the upper and lower bounds of likely future prices, should the pattern hold.*

As shown in Figure 10.7, the short-term bet is that the pattern will continue for a while. Projecting the support and resistance lines beyond today can provide a basis for buy-or-sell decisions. When the price rises to a resistance line, it may be time to sell because the price may go no higher. And when the price falls to a support line, it could be a good time to buy. Or if you're holding the stock, you may want to enter a *stop-loss order,* triggering a sale because that's the lowest price you as an investor are willing to tolerate.

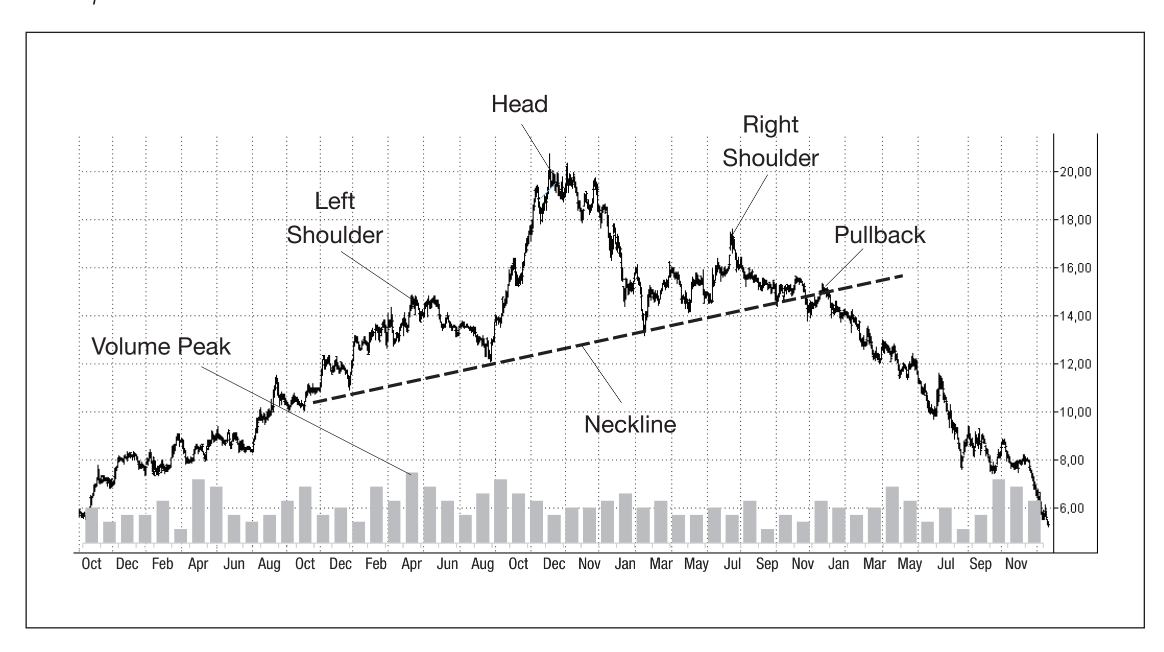
Head and Shoulders Pattern

There are many, many types of chart patterns. To provide an example, I'll talk about one type. And, as I've said, no one knows precisely why they occur. It's just that when a chart pattern appears, some generalizations about the historical behavior of that pattern can make predictions about what will likely happen *within* the pattern.

The pattern I'll use as an example is called *Head and Shoulders,* as shown in Figure 10.8. On the left in the line plot, there's a prominent bump. That's the left shoulder. The peak in the middle is the head. And the bump on the right, which in this case is somewhat higher than the one on the left, is the right shoulder.

This pattern typically occurs over a period of about two months but can take longer, as shown in this example. (We're not dealing with moment-to-moment results here.) These types of patterns are only apparent when you take the longer view.

Figure 10.8 *The early indication that a Head and Shoulders may be forming is an abrupt peak in the volume, which precedes the left shoulder.*



Variations in Patterns

Strictly speaking, this example is a Head and Shoulders–Top, meaning that the head is the high point. There's an inverted variation called Head and Shoulders–Bottom, with the head at the low point. Many chart patterns have their variations, of which inversion is the most common.

Connecting the low points in the pattern creates the *neckline,* which you'll recognize as a support boundary. An upward-trending neckline is an indicator of a *bullish* (increasing-value) market, and likewise downward indicates a *bearish* (losing-value) trend.

The leftmost volume peak, indicating a flurry of trading, marks the left shoulder and the beginning of the pattern. And that's the highest volume of the whole pattern. Again, the analysts have specific rules of thumb for predicting where the pattern will go. To make meaning out of this, it would appear that the abrupt volume peak indicates sudden interest in the stock—but what does it mean when the same pattern occurs in an index? For purposes of making decisions based on technical analysis, the presumed reason for price or volume changes may actually be at odds with conventional wisdom about what the market is currently doing. And that's the main reason why you should exclude these factors from consideration when applying technical analysis.

Following the volume trends, notice that peaks in the bars are mirrored in the line plot of prices. However, the corresponding peaks in the line are slightly out of phase with (lagging) the peaks in the volume bars.

So the appropriate conclusion to draw from a sudden volume peak is that a left shoulder could appear very soon in the price line.

Another rule of thumb, occurring for no apparent reason, is that, within any pattern, the price will tend to return briefly to support lines. This is known as a *pullback.* It's the tendency of a pattern to return to a previous level. A minor pullback is not thought to be significant at all and does not indicate a breakout or start of a new pattern.

And again, the pattern ends when the price breaks out through either the resistance line or the support line—and keeps going.

Bulkowski's Strategy for Head and Shoulders

Renowned technical analyst Tom Bulkowski wrote *Encyclopedia of Chart Patterns*. He claims to have identified some 35,000 patterns, and he has developed strategies for making investment decisions based on the most common ones.

Here's Bulkowski's formula for taking advantage of a Head and Shoulders pattern. It's a good example of how technical analysis makes predictions based on geometry, not market forces:

1. Subtract the neckline price from the Head peak price.
2. Project a future low point by subtracting the result from the neckline at a future point.
3. This result is a minimum target price, which corresponds to the likely end of the pattern.

The investor may use this projected minimum target price as a basis for stop-loss or short-sale orders.

Head and Shoulders—Rules of Thumb

- Pattern typically occurs over two months
- Head and Shoulders peaks are mirrored in Volume plot
- Higher Left Shoulder in Volume indicates strength (bullish trend)
- Neckline connects the troughs of the pattern
- Rising Neckline indicates strength
- Brief Pullback often occurs after price pierces Neckline
- Support level is where prices declined recently
- Support is the short-term low
- Neckline is predictable and can be a trigger point to buy or sell when price pierces the line
- Price tends to decline after Pullback, after pattern (Downward Breakout)
- Bottom may be start of a Rebound pattern (time to buy and hold?)

A *short sale* is a bet that the stock price will go down at a future time. The investor buys an option on the stock at today's price, but then doesn't actually buy the stock until it reaches the stop-loss level, and pays the lower price, making a profit.

Two Different Viewpoints—Fundamental vs. Technical Analysis

A classic work in the field of fundamental analysis is the book *A Random Walk Down Wall Street* by Burton G. Malkiel, a professor of economics at Princeton. Malkiel's philosophy is that markets are generally efficient and future prices are essentially unpredictable, or random. His advice would be to consider the quality of a company's products, the strength of its competition, the competence of its CEO, the current state of its balance sheet. These are traditional, subjective factors, which put emphasis on judgment. And because future results are unpredictable, even when sound judgment is applied to individual stocks, prudent investors should diversify their portfolios, mitigate risks, and pursue strategies that mix *growth* (price appreciation) and *income* (dividends).

$$$$INVESTOR'S TIP$$$$

Mixing technical analysis and fundamental analysis is a pitfall, I think.

A book that provides a deliberate counterpoint to Malkiel is *A Non-Random Walk Down Wall Street* by MIT professor Andrew W. Lo and A. Craig MacKinlay. These authors hold that markets are subject to cycles and group psychology. They believe that chart patterns are not necessarily meaningful but can be useful. Adherents to this point of view are more likely to be faithful practitioners of technical analysis.

Carrying the notion of group psychology a step further, Prof. Lo's research has analyzed historical market behavior and found some support for his theories. He finds that the NASDAQ market, which generally has a higher percentage of individual investors and day traders, is more prone to emotional influences, such as rumors and sensational news events. By contrast, the NYSE, in which there are more institutional investors and programmed trading, tends to be more rational.

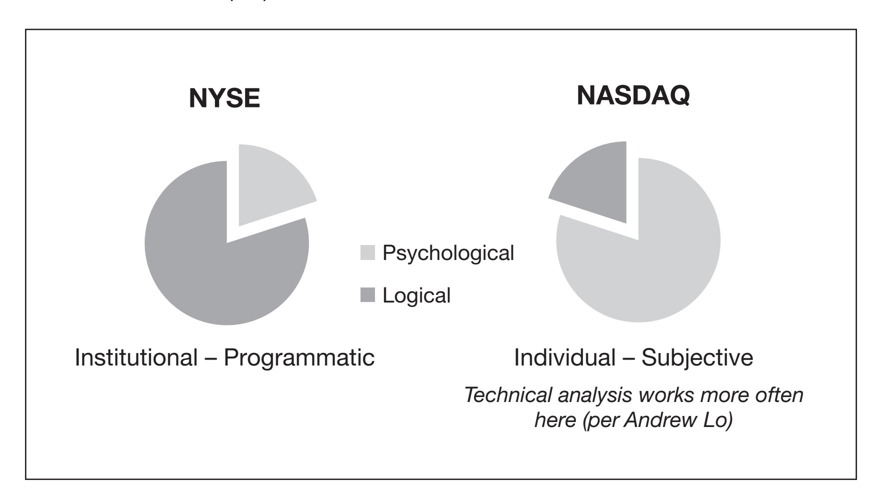
Programmed Trading

Large investors such as banks, investment houses, and pension funds use computer-based algorithms to manage high-volume trading on a moment-to-moment basis during the trading day. Predefined trading policies and transactions involving thousands of shares can be triggered automatically—in milliseconds. In this environment, even the briefest delays can have major financial consequences. As an example, see the discussion on the Flash Crash of 2010 later in this chapter.

Bulkowski goes even further to assert, "Chart patterns are...the footprints of the smart money." This statement would seem to say that patterns mirror the strategies of institutional investors. However, Lo's research suggests the opposite—that is, technical analysis is looking at herd mentality.

Figure 10.9 presents a generalized and notional picture of Prof. Lo's findings. It's a gross calculation, not a precise measurement, but he's saying, in effect, NYSE trades are driven 20% by psychology and 80% by logic, whereas NASDAQ is more like 80% psychology and 20% logic. The

Figure 10.9 *Prof. Andrew Lo's research suggests that market motivators are a mix of psychological (perhaps emotional or sociological) and logical (programmatic) factors and that these motivators may influence different markets in different proportions.*



psychological factors would range between fear and confidence. The logical factors would depend more on metrics, such as quarterly results against projections, interest rates, price-earnings ratios, and dividends.

Interestingly, these diametrically different approaches to analysis seem exactly opposite in their reliance on formula. That is, in looking at company and market performance data, fundamental analysis must be more subjective because it requires interpretation of results. But technical analysis, in parsing the charts of emotional herd-like behavior, is (or perhaps should be) entirely mathematical, ignoring or at least discounting any explanations for the patterns.

Other Types of Patterns

Just to give you an idea of the variety, here's an abbreviated list of other chart patterns. From their names you can tell they come in all different shapes:

- Inverted Head and Shoulders
- Double Top (M-Top)
- Double Bottom (W-Bottom)
- Cup and Handle
- Pennant and Flag
- Bump and Run Reversal (BARR)
- Symmetrical Triangle
- Ascending Triangle
- Doji
- Three Peaks and a Domed House
- Swing Trade
- Diamond Bottoms and Tops
- Scallops
- Islands
- Pipes
- Falling Peaks
- Rising Valleys
- Horns
- Gaps
- Wedges
- Dead Cats!
- Earnings Surprise
- Reverse, Inverted, Double, and Triple Variations

But What Do Chart Patterns *Mean*?

By now you probably know I'm going to dodge this question! In fact, I think it's the wrong question. To apply technical analysis successfully, I think you have to focus on what price levels the math predicts, *not on what's causing them.* The predictions made by technical analysis may be quite clear. That's because those predictions arise from if-then propositions rooted entirely in statistical probabilities within the pattern, not in external circumstances or influences.

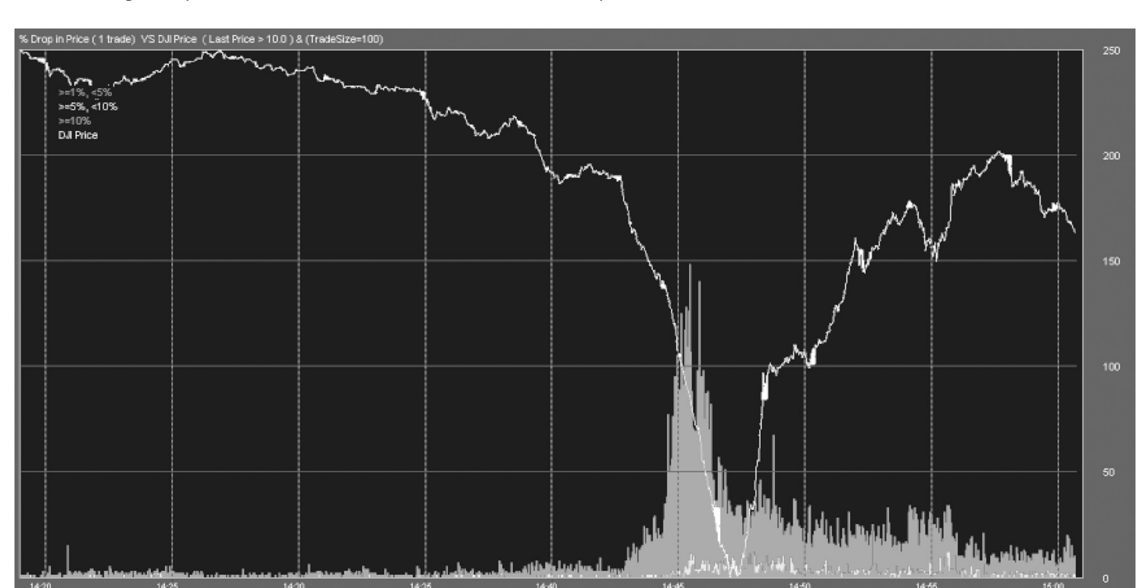
Technical analysis gives you a limited set of options when other financial gurus may be debating the fundamentals and coming to widely different conclusions. In fact, it might seem perfectly reasonable to say, "Technical analysis is a tool." But taking this approach suggests that you will be supplementing it with other methods of analysis. By doing so, you'll be reading meaning into the chart patterns and, I suggest, confusing yourself.

That said, here's the only piece of specific investment advice you'll get from me: Never plan a trading tactic without a corresponding *stop-loss order,* a standing instruction to your broker to sell a particular stock at a specified low point. Pattern analysis should give you a fair idea of the support levels. Pick a price below which you will not go, and plan to sell the stock if it ever goes there. Thus you will mitigate risks such as making a wrong prediction, misidentifying the pattern, or entering the pattern at the wrong point.

Pattern Recognition—the Scary Part

On May 6, 2010 the NYSE DJIA lost about 600 points within minutes. At the worst of it, prices dropped 200 points in the astonishingly short interval of about 10 seconds (Figure 10.10). This event became known as the "Flash Crash." The market soon recovered, and the Securities and Exchange Commission (SEC) launched an investigation into its causes. Ultimately, the SEC blamed a single mistaken entry of about 75,000 contracts, which had a ripple effect through concurrent trades because it was so unexpectedly large.

Figure 10.10 *This chart shows the worst of the Flash Crash of 2010, showing the average price of the DJIA index losing 200 points within about 10 seconds. (Courtesy Nanex)*



Remember—and this is an important distinction—the Flash Crash happened in a micro time period—during a small fraction of a single trading day. Therefore, it is not the type of pattern studied by technical analysis, which normally spans weeks and months.

But there's a basic similarity in that the root causes of the patterns are unknown, perhaps unknowable. According to official news releases at the time, the Flash Crash mistake was made during a manual entry—keystroke error—not by a glitch in a computerized trading program. But the immediate response to the error was highly automated—as programmed trading algorithms kicked in to trigger a precipitous price drop.

Some observers of the Flash Crash speculated that the triggering "error" may in fact have been deliberate. One such technique that may aim to manipulate programmed trading is called *quote stuffing.* In this scheme, a computer program makes a large number of requests for a price quotation on a particular stock in a very short period of time (on

the order of milliseconds). Other programs sense this activity, and their logic may conclude that a large buy order is coming. So those programs may execute a flurry of their own buys. The result may be that the quote stuffer ran up the price of the stock without actually buying any of it.

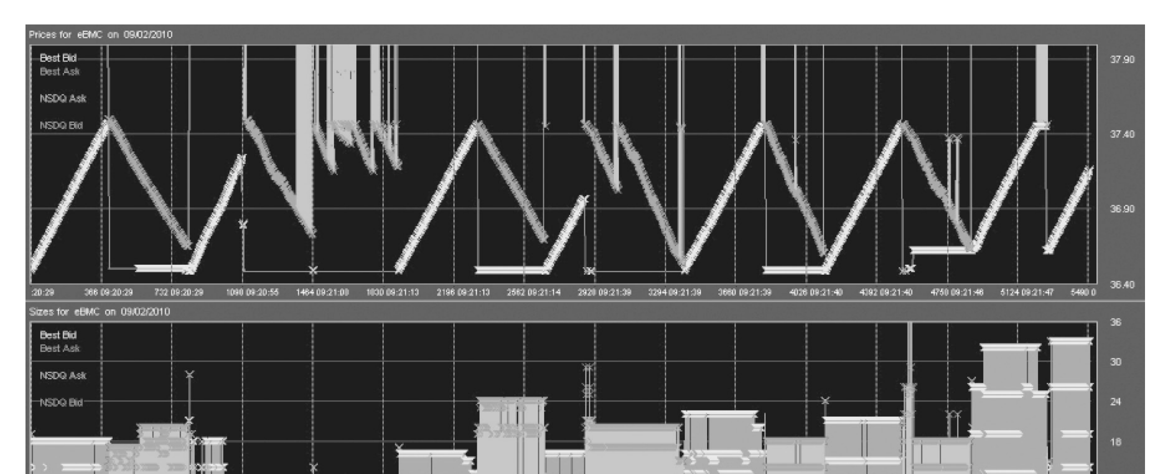
In July of 2013, the Financial Industry Regulatory Authority (FINRA), an industry watchdog organization, announced it was looking into programmed trading practices, with a view toward preventing overreaction of the market as a result of either errors or attempts to game the system.

Then, in the spring of 2015, a financial trader was arrested in the UK and accused of entering and withdrawing thousands of orders over a span of hundreds of trading days in ongoing attempts to manipulate futures markets. This practice is called *spoofing*. The Flash Crash of 2010 is also being blamed on him, as he allegedly placed large orders repeatedly over several hours. (In 2016 the perpetrator plead guilty and began to work with authorities to prevent such abuses.)

Soon after the arrest, *The Wall Street Journal* ran an article, "'Flash Crash' Overhaul Is Snarled in Red Tape." The title pretty much says it all. A consortium of regulators and financial exchanges has been trying to implement real-time monitoring of stock and options orders under a set of standards they are calling the *Consolidated Audit Trail (CAT).* Work on CAT has been underway as far back as 2012, but its goals have so far been elusive. This much is certain: The experts now firmly believe that the Flash Crash was triggered deliberately. And they fully appreciate that the speed and complexity of automated systems compound the problems of detection and diagnosis. The CAT should provide detailed evidence for diagnosis after the fact. But it still might not be possible to detect a malicious action in time to prevent its consequences. (At the time of editing the fourth edition of this book in 2018, CAT was in the early stages of implementation.)

The accused man in the Flash Crash incident was a trader, not an outsider. But there's at least one other major source of worry—that short-duration anomalies in market results may be occurring as a result of hacking attempts, not necessarily by bona fide traders. The attacks may be malicious and done for a variety of purposes, not least of which may

Figure 10.11 *Highly regular patterns appearing in real-time feeds from financial markets are so baffling that analysts call them* crop circles. *(Courtesy Nanex)*



be to simply prove it can be done. Nanex, a firm that provides market-tracking software and services to brokers, finds odd patterns in its charts from time to time. Like the Flash Crash, these are generally very short in duration but also highly regular. (It's not noise.) The shapes are so bizarre that Nanex calls them *Crop Circles,* and the similarity doesn't end there because the causes are equally baffling (Figure 10.11).

Here's the important difference—technical analysis is generally not subject to malicious mischief or gaming the market. Its timeframes are sufficiently long to aggregate results so that short-term variations do not affect further results. (Minor pullbacks are not considered significant, for example.) It is therefore truly the study of the long-term trends of people acting *en masse* rather than patterns arising from anomalies generated by a single trader or program.

There is, however, a point at which technical analysis and programmed trading may converge. Consider that programmed trading algorithms are proprietary to financial institutions. So, the bases of their trading decisions are not public information. This is my considered opinion, and not something for which I can offer empirical proof. But in the future, or

possibly even now, programmed trading may incorporate the principles of technical analysis to project trends. The programs may then execute based on those trends. To the extent this happens, the characteristics of the patterns will feed back on the trading processes. And the accuracy of technical analysis could become a self-fulfilling prophecy, even in markets that seem to be driven more by logic than emotion.

Applying (or Misapplying?) Fibonacci

If you are motivated to investigate the subject of technical analysis further, your searches will no doubt turn up references to Fibonacci. The Italian mathematician Leonardo Pisano Bigollo, better known as Fibonacci, wrote about a special sequence of numbers in which each number is the sum of the previous two. This mathematical formula (probably with its roots in India) describes certain types of organic growth, such as population increase, the spiral of a seashell, or the flowering of plants:

$$Fn = Fn-1 + Fn-2$$

That is, the next number in the series (F_n) is equal to the sum of the two preceding numbers (given the first two numbers, 0 and 1). The resulting progression is:

0, 1, 1, 2, 3, 5, 8, 13, 21, 55, 89, 144...

$$$$INVESTOR'S TIP$$$$

Is there a "fib" in Fibonacci? Some aspects of this analysts' rule of thumb remind me of the fudging that golfers do when they factor their handicaps into their actual scores.

Now, there are certain wizards of finance who tout the Fibonacci series as the key to understanding some financial market patterns. However, I find their logic flawed (or perhaps I misunderstand their approach).

As I understand it, these fans of Fibonacci observe that, after 2 in the series above, every succeeding number is about 62% higher. (This rule of thumb is itself a rather crude approximation.) They take this to be the maximum amount of gain or loss when projecting the bounds of an emerging pattern. For the minimum amount of gain or loss, they subtract this percentage from 100% (for reasons that are opaque to me), to arrive at 38%. The average amount of gain or loss, they reason, is the midpoint between these two bounds (Figure 10.12).

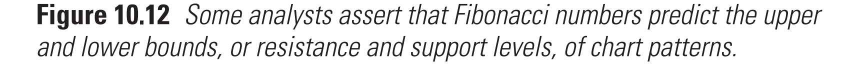
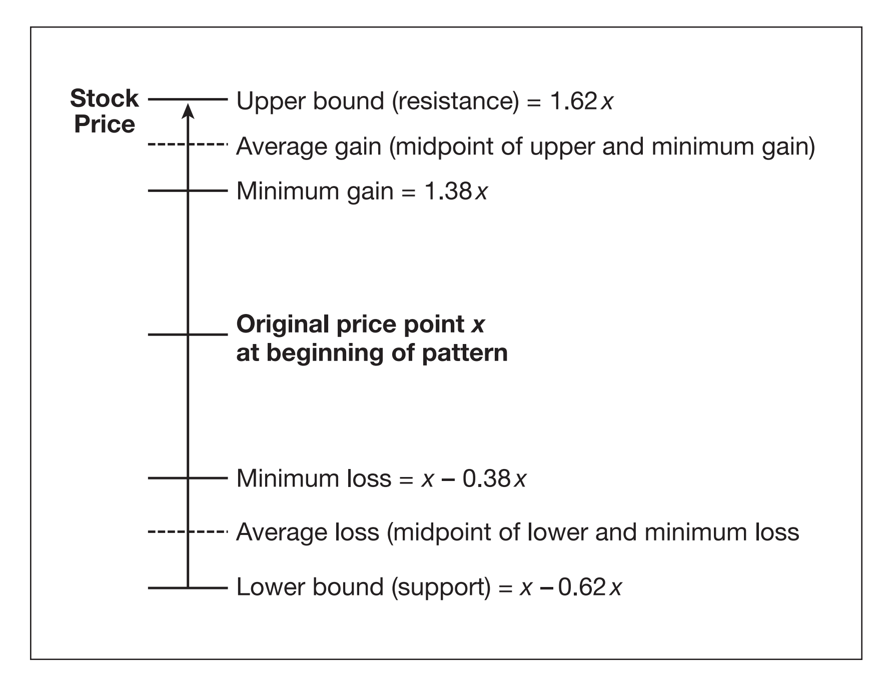


Figure 10.12 *Some analysts assert that Fibonacci numbers predict the upper and lower bounds, or resistance and support levels, of chart patterns.*



One problem with this theory is that a Fibonacci series presumably is an infinite progression, which has no midpoint. So the notion of a midpoint is extraneous to any correspondence between the formula and financial-market events. It's rather like the golfer's handicap discussed in Chapter 1, a correction factor applied to bring a theory into agreement with observations.

An appealing aspect of relating Fibonacci numbers to the markets is the implication that both somehow obey the forces of nature. This concept aligns with the notions of an "unseen hand" guiding the market and the "wisdom of the market" transcending that of individual experts. While it is possible that the benchmarks of plus or minus 62 or 38% could be useful in predicting price levels, any such success would appear to be only loosely related to Fibonacci's formula.

Appreciating Technical Analysis

This chapter is just a toe-dip into the ocean of technical analysis. In the jargon of the college curriculum, it's an "appreciation course," a survey of the topic so you can see what it covers and what some of its issues are. However brief as this discussion is, there are a few important takeaways:

- It is an automatic human reflex to find meanings in patterns, including financial market charts.
- Because of this irresistible tendency, even professional financial analysts will form "hunches" about the reasons for market fluctuations. These interpretations *can* be a proper focus for fundamental analysis. But the reasons behind price changes have nothing at all to do with the *statistical* predictions that proceed from technical analysis.
- Technical analysis does not explain, nor does it attempt to explain, the reasons behind market fluctuations. So be careful about mixing fundamental analysis and technical analysis in making financial decisions.

Unlike some other chart types I describe in this book, I don't think financial market charts are inherently misleading. The main pitfall in applying technical analysis is to construe real-world motivations in the patterns. However, as the next chapter demonstrates, some methods of displaying financial charts can invite misinterpretation and faulty conclusions.

11

Viewports and Dashboards

Not Quite As Bad As a Blurry Screen

THERE are cases when a chart may not be misleading in itself, but the method of display introduces some distortion. Such distortion can be particularly troublesome when multiple charts are shown on a page or screen, inviting comparisons among them.

A *viewport* is a chart-display window on a Web page or application screen. A viewport is typically a fixed size, a limited amount of screen "real estate." The software that populates the window is not looking for truth in chart presentation. It is looking to fill the window most attractively—or most efficiently. For example, if selecting the *x*-axis scale range for a chart would make it look all squashed up at the bottom, the result will look like a mistake. So the software will expand the scale to fill the window.

Put another way, charts are usually scaled to fit the viewport, not the viewport to the requirements of the chart. And that's why viewports can be troublesome.

So, a pitfall of interpreting *xy* charts has to do with scaling of the axes. I cover some of these issues in Chapter 6. This chapter extends that discussion to the problems that arise when it's necessary to fit one or more charts within a fixed-sized area.

A *dashboard display* is an arrangement of multiple viewports. Problems of scaling within viewports are compounded when they are presented side-by-side.

A related issue is whether charts that are presented alongside each other are truly comparable. A comparison is, after all, a way of deriving meaning by looking for similarities and differences between patterns.

Chart Comparisons

$$$$INVESTOR'S TIP$$$$

My Golden Rule of charting applies here: *Do as you would want to be done to.* Once you see how misleading financial charts can be, you'll feel you are not only a better interpreter of charts but also a better maker of charts in your business reports.

Here's how that auto-scaling can get tricky, especially when chart windows are shown side-by-side.

Figure 11.1 shows two *xy* charts adjacent to one another. Each of the charts has been scaled to fit the window and not in a way that makes the chart scales proportional. As a result the charts are *not* visually comparable. It looks as though the two plots have the same slope, but they don't. As just mentioned, the software has apparently stretched the scale of the chart on the right to make it fill the viewport.

In fact, Acme Corp has experienced a 100% increase, and Widget just 15%, but within widely different price ranges. The Acme Corp stock rose from $16 to $33, and Widget went from $96 to 115. You could plot both of these on a scale from $15–115, but because of the expanded scale—with so much empty space in the middle—both plots would be flattened so much you wouldn't detect any increase at all.

Showing the scales with different but proportional ranges results in the pair of charts shown in Figure 11.2. In this case, the ratio between the upper bound of the Acme chart and its lower bound is:

$35 / $15 = 2.33

And the ratio between the Widget upper and lower bounds is nearly the same:

$215 / $95 = 2.26

Scaling the two charts by the same proportions makes the slopes of the two plots proportional, as well—and therefore comparable.

Note that if you're using an application such as PowerPoint or Excel to generate these charts, you'll have to do this axis scaling manually. That

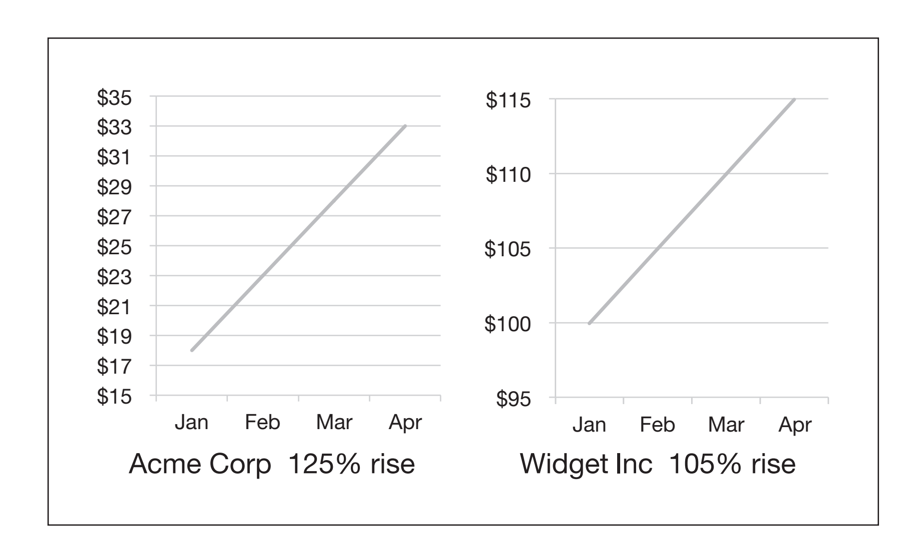


Figure 11.1 *The chart-generating application has scaled these two charts independently, each so it fills its viewport as completely and attractively as possible. However, the sizing of the y-axis scales to fit the windows is not proportional, making comparisons between the charts difficult.*

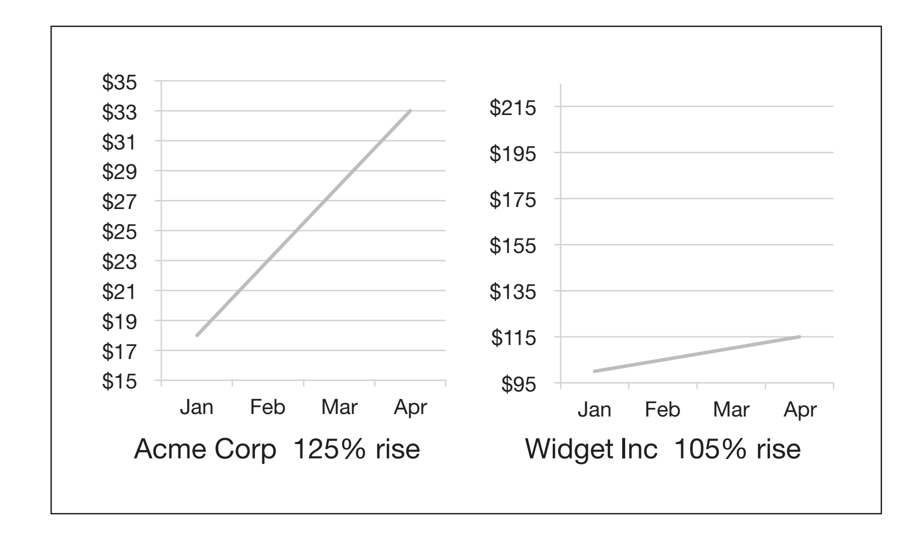


Figure 11.2 *Scaling the Widget chart in the previous figure to be proportional to Acme results in slopes that are comparable.*

is, if you simply plot the Widget chart from the data range $96–115, the software will probably generate a *y*-axis scale ranging from $95–120. You'll have to do the calculation to find out how high the upper bound should go to be proportional to the companion chart, then go in and specify that upper bound in your chart.

Vectors vs. Bitmaps

From a technical standpoint, scale distortion is much more likely when the chart being displayed is generated as *vector graphics,* as in Windows Metafiles (.wmf). That's because the application software can resize and render the geometrical objects in such files "on the fly" to fit the display. By contrast, a bitmap, such as a JPEG (.jpg) file, is pre-rendered. It is simply an array of pixels. Its scaling therefore can't be distorted unless the entire image is squeezed or squashed. All that said, from the programmer's perspective, vector graphics are preferable to bitmaps precisely because their proportions and resolution are readily adjustable to fit a variety of screen formats.

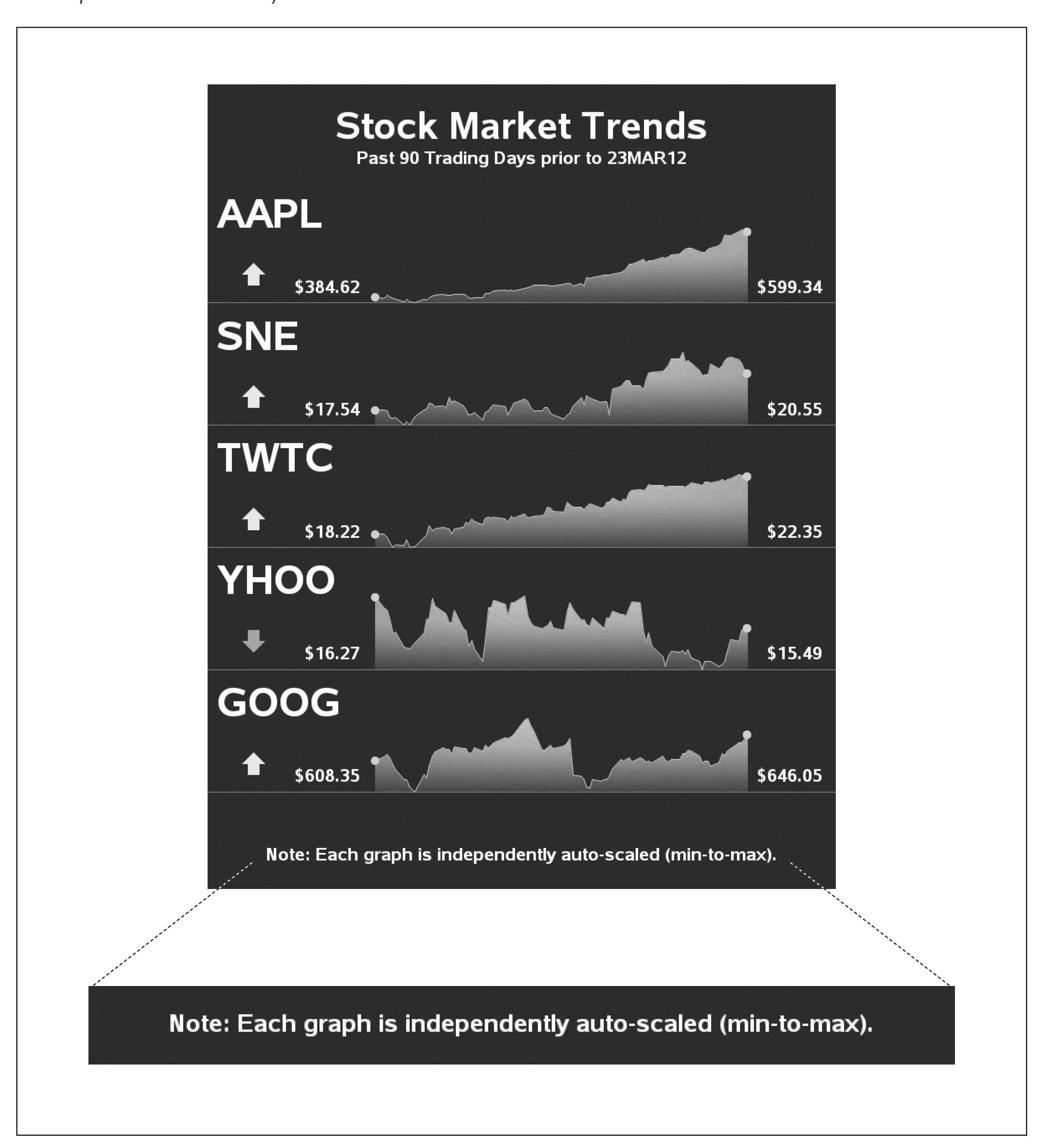
Smartphone Example

Figure 11.3 has an example of greatly mismatched chart scales in a comparison that is inherently misleading. Here are results for five different stocks. (The time period shown is Q1 of 2012 so these obviously aren't today's prices.) Over the 90-day period, Apple (AAPL) stock rose from $384.62 to $599.34. During that same period, Sony (SNE) went from $17.54 to $20.55, TW Telecom (TWTC) from $18.22 to $22.35, Yahoo (YHOO) from $16.27 to $15.49, and Google (GOOG) from $608.35 to $646.05.

Because these stocks fluctuate over such widely different ranges, there is no practical way to plot them all against the same scale—and still fit them on the face of the smartphone. So here, each stock gets the same amount of screen real estate, and the plot of prices is fit inside to fill it.

The chart makers—in this case the programmers who created the stock price app—are aware that this is inherently misleading. The cautionary note at the bottom reads, "Each graph is independently auto-scaled (min-to-max)." So if, for example, you're tempted to make some conclusions about which of these stocks is the most volatile—don't!

Figure 11.3 *Comparisons among these five stock-price plots, such as judging relative volatility, aren't meaningful because the* y*-axis scales cover widely different ranges. The app designers are obviously aware of this problem because they've inserted a note of caution at the bottom.*



Dashboard Displays

Dashboard displays bear some similarity to instrument and control panels on automobiles and aircraft. They typically show multiple charts and tables of results on the same page. For years, the dashboard display has been the preferred format for so-called *executive information systems,* which present summary data from data warehouses. In such applications, users expect the dashboard display to be always in the same format—with results from each business entity, department, or functional area appearing in the same location day after day. This layout presumably makes at-a-glance status checking easy because the user can concentrate on variations and exceptions.

You can probably identify the pitfall here. A dashboard display is essentially an arrangement of separate viewports. And as previously pointed out, the application's graphics software will tend to scale

Designing a Dashboard

My recommendations for design of dashboard displays would be:

- The dashboard should be laid out for clarity—viewports spaced evenly, without clutter. Charts should be large enough so that labels are no smaller than 8pt type.
- Viewports for different entities should appear in consistent location from day to day. Chart scaling within viewports should be as consistent as possible from day to day. Ideally, scale breaks should be used and clearly marked, but this is seldom done because chart apps don't work that way.
- All 3D chart formats should be avoided, for reasons outlined in "Pies, Clocks, and Compasses" in Chapter 2 (pie charts) "Tricks of Perspective" in Chapter 4 (bar charts). For example, it's difficult to visually estimate the true height of 3D bars against the chart scales. This problem is compounded when 3D bar charts are presented side-by-side.

each chart to fill the available space in each viewport. So a day-to-day variation that's exceptionally large or small may be effectively hidden by the adjustment in chart scales that is necessary just to fit the chart within the viewport.

Figure 11.4 is an example of a dashboard display that does a nice job of avoiding these inherent problems.

Figure 11.4 *This dashboard display is an example of good practice: 1) The arrangement of viewports is symmetrical and uncluttered. 2) Viewports have consistent and proportional scales. 3) The charts don't use 3D bars.*



12

Words for Your Charts

Or, Is There Any Truth in Labeling?

INADVERTENTLY **leaving off helpful labels** on a chart is the quickest way to perjure yourself. In Chapter 1, I invoke the memory of my high-school chemistry teacher, Mr. McNeill, who punished me mercilessly (on paper) for omitting labels from numeric answers. In this chapter, I'll show you how data labels can clarify or confuse the interpretation of charts, and I'll expand the discussion considerably to include guidelines for all kinds of explanatory text in presentations—including the text of your speech or report.



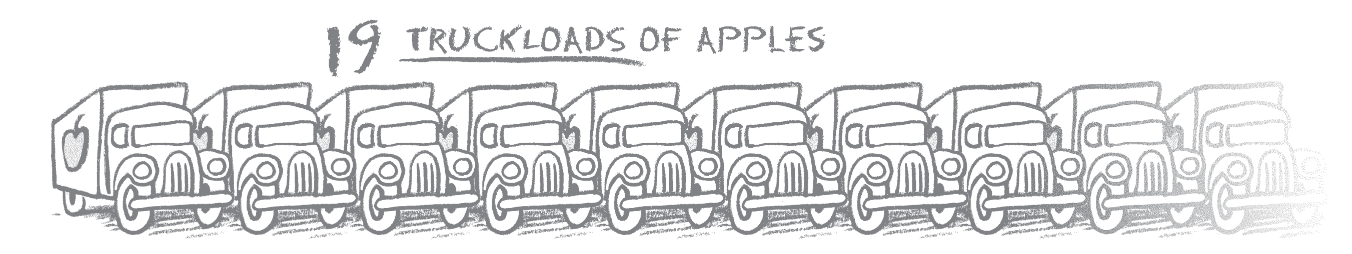
Another of my teachers, French philosopher Roland Barthes, said that, syntactically, a book is a long sentence spoken by its author. It might be productive to think of your report as a long caption for your charts. If you say something that doesn't have a picture to go with it, come up with a new picture—or get rid of the material.

Did I Give You Data or Information?

Labels can mean the difference between having raw data and having useful information. To recap Chapter 1, the categories of things on your charts can be indicated with labels, which are as necessary to reliable data as the numbers(or *values).* As a rule, wherever you see a number, you should find a label nearby.

19 APPLES!

Often a label must have two parts: a unit of measure, and a description of the thing being measured.



$$$$INVESTOR'S TIP$$$$

Some short-sighted programmers ignore the importance and usefulness of specifying delimiters in the data fields of online forms. Delimiters help users "chunk" data for ease of reference and retention and can help prevent keystroke errors on data entry.

Labels on data can sometimes be hidden or implied. Certain conventions for displaying results are shortcuts to labeling data. One type of shortcut involves using a specific style of punctuation to subdivide a set of numbers. To pick an example I've used previously, the following set of numbers has no obvious meaning:

8005551212

But if I set them off by using a particular style of punctuation, you will have some valuable information:

(800) 555-1212

Aided by the helpful parentheses and hyphen, you might recognize this set of digits as a North American telephone number. In computer jargon, the punctuation marks are *delimiters*, and they mark the divisions between information *fields*. (For an explanation of fields in data records, see Chapter 9.) The hidden labels are the implied names of these fields, which you know by the convention for composing phone numbers:

Area Code	Exchange	Number
800	555	1212



It is a liar's trick to mislabel even the data that has hidden labels by using delimiting conventions that are either unfamiliar to the audience, or incorrect in the context of the presentation.

Here's an example of another set of delimiting conventions, a date written as a set of numbers:

09062049

In the United States, an audience might expect to see the date delimited as follows:

09-06-2049

or

09/06/2049

The implied fields are:

Month	Day	Year
9	6	2049

For an audience in the United States, then, the date will be interpreted as:

SEPTEMBER 6, 2049

But if I use another style of delimiter, you may get a different clue about the interpretation of the date:

09.06.2049

The periods are not typically used in the United States for showing dates, but this is a common convention in Europe. And there, the implied labels are different:

Day	Month	Year
9	6	2049

In Europe, the correct interpretation of the same set of digits would be:

JUNE 9, 2049



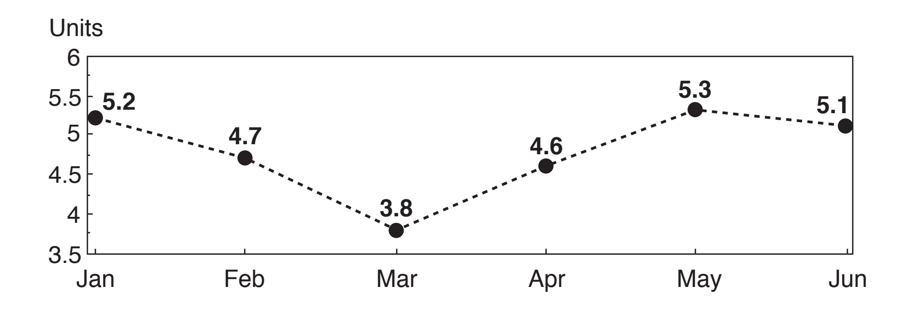
The example of different date conventions might seem trivial, but it's symptomatic of more sophisticated liar's tricks. For example, when the presenter speaks of results for "the current year," is the reference to the calendar or to the fiscal year? Any confusion about which accounting period is being discussed could have major financial implications, in which case a misunderstanding about date format might not be trivial at all. Any company that does business internationally must establish consistent standards for such simple rules as formats for monetary amounts and dates which appear in reports and presentations.

The Golden Rules of Data Labeling

There are generally two types of labels on charts: those on data points and those on axes. Here are some guidelines for each type.

Labeling Data Points

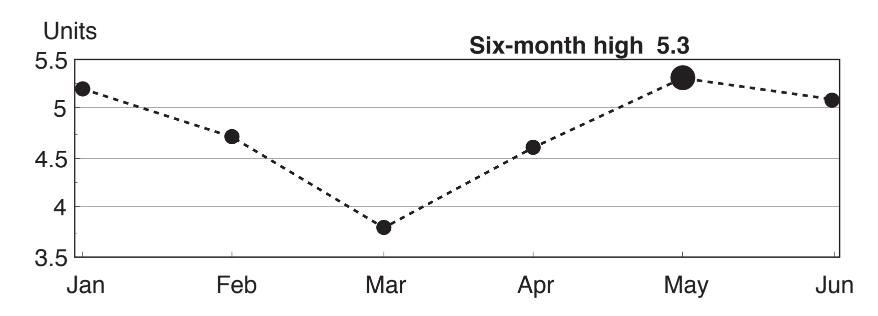
A basic tool of chart designers is to label the values of data points on a line plot or set of bars:



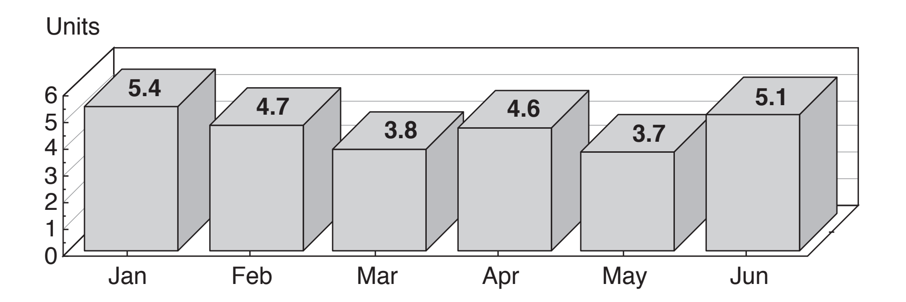
You needn't include data labels on all your charts. In fact, too many labels can clutter the design and actually distract from the

interpretation of the graph. But there are circumstances when labeling should be mandatory. In general, you should use data labels to:

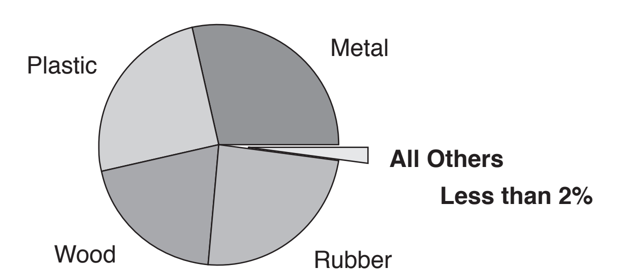
- Identify a specific data point you want to highlight in your presentation, such as an exceptionally high or low value, or an average:



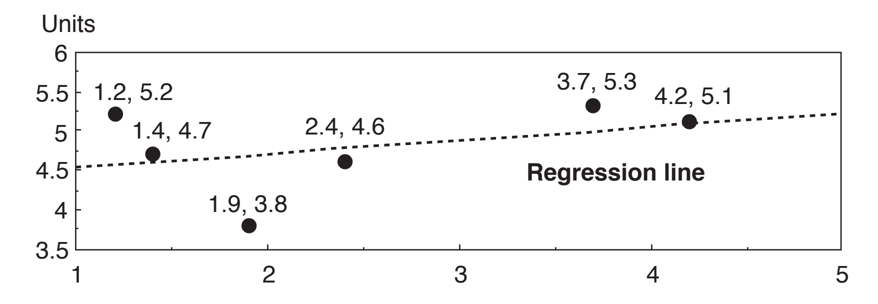
- Make sure that the audience interprets the correct value on plots which are difficult to read, such as estimating the height of 3D bars in relation to the chart scale:



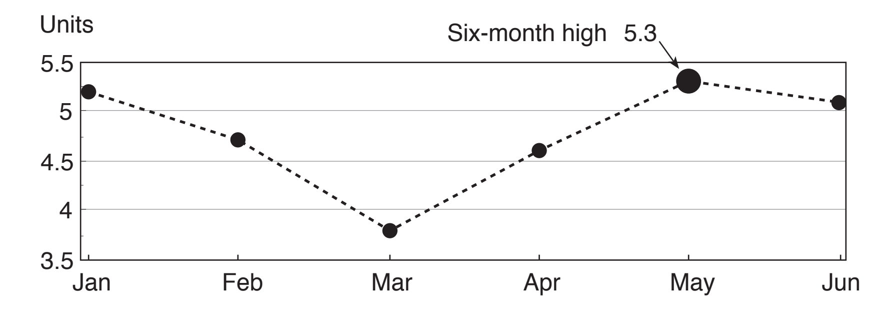
- Show the actual value for a small bar segment or pie slice that had to be made bigger just to be visible:



- In an *xy* scatter chart, identify actual values from which trend lines were calculated (such as by extrapolation, interpolation, or regression):



- If the design does not permit locating the label close to the data point, use a *callout line* to direct the eye:





As I emphasize in Chapter 2, you should *not* show labels of actual values on pie charts or in any chart format that shows proportional relationships. If the purpose of the chart is to show proportions, use percentage labels instead. If you must show actual values, do it on a separate chart—perhaps formatted as a vertical bar graph.

Options for Labeling Chart Axes

There are two elements of labeling the axes of *xy* charts: axis titles and scale division labels.

Axis Titles

An axis title simply identifies what type of measurement the axis represents. Here are the rules:

- Include both the type of measurement and the unit of measure in an axis title:

SALES ($)

- You can omit an axis title, such as Weekday, only if it would be obvious from the scale labels:

MON TUE WED THU FRI

- If a scaling factor, or multiplier, is needed to interpret the scale labels properly, it must be included either in the axis title or in a footnote to the chart.

SALES ($ MILLIONS)

*SALES DOLLARS IN MILLIONS

- If an *xy* chart has dual-*y* axes, you must show an axis title for each. Ideally, the axis titles and legend entries should be color-coded to their respective data series. (See Figure 12.1.)

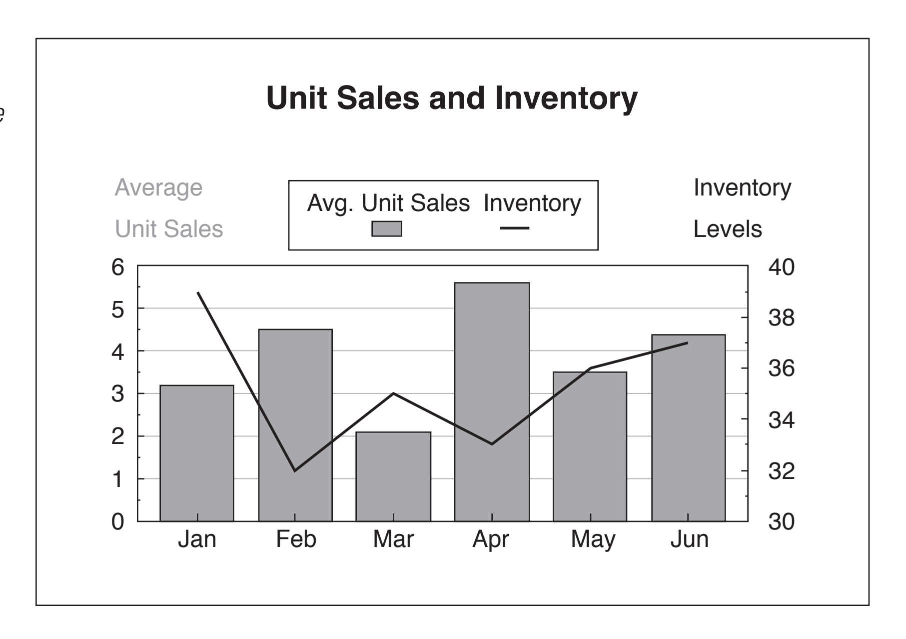


Figure 12.1 *Charts with dual-y axes should have y-axis titles that are color-coded to legend entries and plots.*

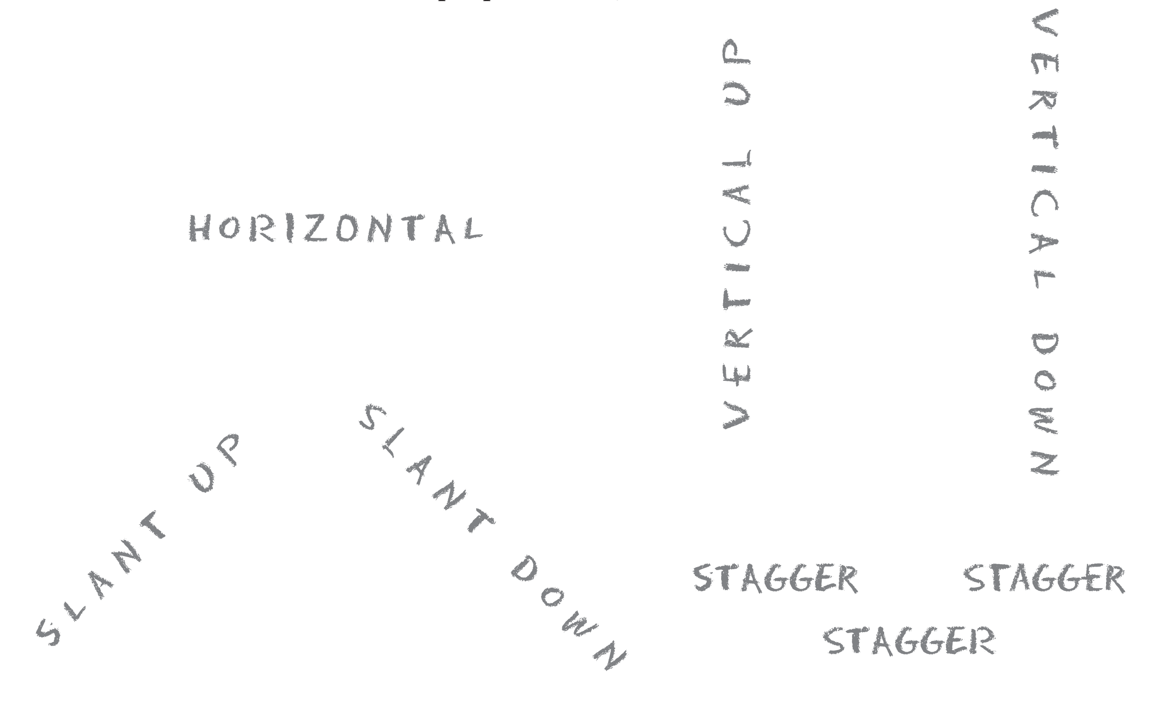
Scale Labels

The honest trick with scale labels is to make them as large and readable as possible. (You'll find some guidelines on choosing type sizes later in this chapter.) Besides increasing the type size, there are two other techniques you can use to enhance readability of scale labels:

- Use a *skip factor* to omit labels and regular intervals from a sequence. For example, if the skip factor were 5, every fifth label would be printed on the scale:

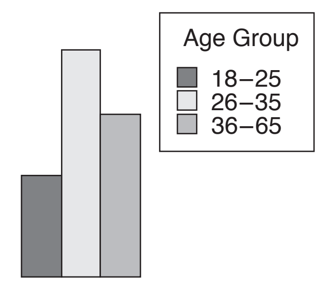


- If after using a skip factor the labels still seem cluttered, adjust their orientation. The normal orientation of labels is horizontal. Some of the alternatives are: vertical up, vertical down, slant up, slant down, and stagger. (Vertical up might still be difficult for some people to read.)

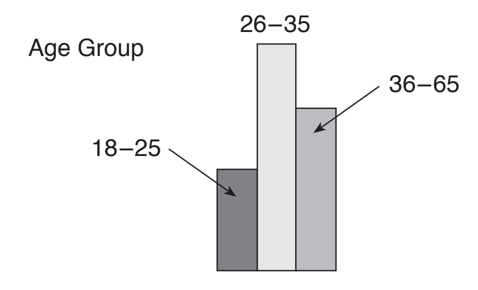


Composing Legends and Data Tables

A chart *legend* correlates the color or pattern of a plot to the name of its data series:



Another term for *legend* is *key*. If there aren't many data series in a chart, an alternative to providing a key is simply to label the data plots:





It's a liar's trick to use color codes that are too similar, making it difficult to correlate the names of data series with their plots.

An alternative to a plain legend is a data table, which shows the numeric values for each data series, usually beneath the chart area. It is customary to show a legend or a data table, but not both. The data table usually incorporates the legend:

Age Grp.	Movies	Web	Television	Newspaper
18 - 25	35	40	23	2
26 - 35	30	42	24	4
36 - 65	26	39	30	5



On the TV quiz show *Wheel of Fortune*, contestants have to buy each vowel to solve word puzzles. Chart liars must think the same way: They use labels so sparingly that you'd think they're being charged a high price for each letter!

Some Notes on Footnotes

$$$$INVESTOR'S TIP$$$$

Footnotes in financial reports need special attention to form and consistency. See Chapter 13.

In a chart, footnotes can be used to insert citations of sources of data and other explanatory material, as shown in Figure 12.2.

If there is a report narrative accompanying a printed chart, the form of a footnote usually requires a superscripted number after the sentence in the text to which the note refers.[1] In the most common format, the numbered footnote appears at the bottom of the same page.

As far as I know, there are no formal rules of style for footnotes in visual presentations. However, citation of information sources can be a serious matter. In general, you can be guided by the rules for printed

1. Lengthy notes also can be continued on the following pages of a report. However, breaking footnotes between pages is not only troublesome for text composition, but it can also be confusing to the reader.

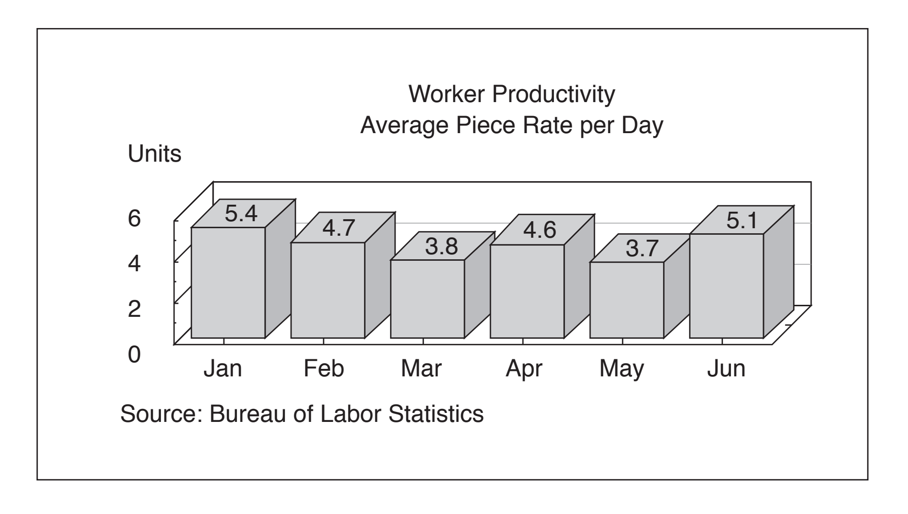


Figure 12.2 *Use a footnote to disclose the source of information for a chart or the grantor of permission for reproduction.*

publications, where the form of the citation depends on the type of source—whether from a textbook or from a scholarly journal, for example. For the form of footnote references and citation of sources, see *The Chicago Manual of Style* (University of Chicago Press).

It's Not Just Polite, It's the Law

An important use of footnotes in both reports and presentations is to give credit to the organizations or individuals who hold intellectual property rights to the material. Again, the rules are more formal in print, but you should be aware of some of the legalities involved.



The liar's trick with footnotes is pretty obvious: They don't use them because they either are not willing to disclose the source of their material—or they have no authoritative source at all!

Requirements for obtaining permission to reproduce quoted material and drawings are covered by national and international copyright laws. Opinions differ on the doctrine of *fair use,* or reproduction of small excerpts from a longer work without explicit permission. If you get formal permission, you should set the *credit line*, or the citation of the source, exactly as specified by the permission grantor. If you use a short quotation under fair use, the source should still be cited.

Credit lines for photographs and illustrations are generally set in the figure captions. Permissions to quote may be included on the copyright page of a long report, which in a book should be on the back of the title page. Permissions might also be included in a separate section of acknowledgments. In more formal documents, such as scholarly articles and books, citations are included in footnotes at the bottom of the page or numbered by chapter in an appendix titled "Notes" or "References." (Be aware that a bibliography is not a list of specific citations, but a separate listing of research sources.)

Related to the subject of permissions is the question of plagiarism. Don't assume that you can get around the need to obtain permission simply by paraphrasing, or rewording, a quotation or by redrawing a chart or illustration. Copyright law uses the standard of *substantial similarity*

to define plagiarism—and leaves it to the courts to decide what that means in each case.

Some material is so old that it is out of copyright, or in the *public domain.* This material might not require permission for quotation. But be careful! For example, no one holds a copyright on the words of Shakespeare. However, the *design and typesetting* of his words in a new book edition might be protected by copyright law—so you can't just go reproducing the pages. Also protected would be a new translation into another language, scholarly comments or references included in the text, illustrations, and so on. Likewise, no one has the rights to Leonardo da Vinci's drawings. However, a museum typically would restrict public access to the originals, and a person who was granted permission to photograph those drawings might claim proprietary interest in the right to reproduce the photographs.

Clip art is predrawn art sold specifically to be incorporated into your presentations. However, publication rights vary among vendors. If you plan to offer a report for sale to the public, consult the language of the license agreement that came with the clip art library.



Don't look here for an interpretation of the law. I'm not a lawyer, but I reserve the right to play one on TV.

Readability Is Fundamental!

It's fine to talk about making chart labels readable, but you can't do that without some specifics on type composition. Here is an overview.

What Type of Type Is a Font?

To a professional typesetter, a *font* is a particular design of lettering in a specific size. The design is called a *typeface* (or just *face*).

Type size is usually measured in *points.* There are 72 points to an inch. (The exact specification used in typesetting is 72.254 points to the inch, but who cares?)

Figure 12.3 *The character set of the font Arial*

	!	"	#	$	%	&	'	(	)	*	+	,	-	.	/	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	□
□	□	‚	ƒ	„	…	†	‡	ˆ	‰	Š	‹	Œ	□	□	□	□	‘	’	“	”	▪	–	—	˜	™	š	›	œ	□	□	Ÿ
	¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

All of the characters in the Arial font are shown in Figure 12.3. This font is similar to Helvetica, which is probably the most widely used set of letters in the world.

All Upstanding Fonts Have Character

A font generally includes all of the letters of the alphabet in capitals (A–Z) and in lowercase (a–z), as well as the numeric digits 0–9, punctuation marks, special characters such as the dollar sign ($), and perhaps some characters used in other languages (usually referred to

The Short Story of the Points and Picas

Standard single-spacing in letters and reports is 6 lines to the inch. This is also the definition of the unit of measurement called the *pica.* In conventional typesetting, the point unit is used with type, and the pica is used for measurements on a page, such as leading and the widths of lines and columns.

However, on most personal computers, *all* numeric settings—including type size, letterspacing (space between letters in a word) and leading (vertical space between lines)—are in points. Vertical spacing is usually controlled as a multiple of the height of a line, which depends on the point size of the type: The bigger the type, the more space between the lines. So, single-spacing will produce 6 lines per inch *only if the type size is 12 points* (6 lines/inch × 12 pt = 72 points/inch). It should come as no surprise, then, that 12-point typewriter type was called Pica: Each character is one pica tall.

as *international characters*). All of the characters in a font constitute its *character set.*

Fonts differ in the number and variety of characters that they contain. For example, some special-purpose fonts have only capital letters. Most English-language fonts include the British pound sign (£), but many of them do not have the copyright symbol (©).

When Is a Font Just a Font?

Except to graphic arts professionals, the term *font* usually means just *typeface,* without reference to size. That's because most computer fonts are *scalable,* or can be adjusted continuously over a wide range of point sizes, typically from 1 (very small) to 999 (very big).

In software manuals and in books like this one, you will sometimes see the term *font* used rather loosely to refer to the whole set of appearance options, or *attributes,* that are available for a particular typeface.

Attributes Are the Spice of Type

Attributes are ways of varying the appearance of a typeface. Font attributes can be applied to individual characters, or letters—or to groups of letters, such as words or even entire paragraphs or documents. Examples of font attributes include bold, italic, underline, double underline, strikethrough, color, hidden, small caps, all caps, spacing, superscript, subscript, and drop shadow. Most of these attributes are shown in Figure 12.4.



Knowing about font attributes is particularly useful when you are composing charts for reproduction in black-and-white. Attributes such as bold and italic can be used instead of color to highlight labels and emphasize text. (By convention, underlining was once used as a substitute for italics in Roman typewriter text. But these days, avoid underlining to prevent confusion with hyperlinks.)

Figure 12.4 *Font attributes provide many ways to alter the appearance—and the impression given by—a single typeface. The font used in this example is Times New Roman.*

Roman
Bold
Italic
Single underline
Single underline, words only
Double underline
Hidden doesn't print!
~~Strikethrough~~
SMALL CAPS
ALL CAPS

Shopping for Just the Right Text Font

$$$$INVESTOR'S TIP$$$$
De facto standards for government documents such as proposals and reports is 12-pt Times New Roman for text and 10-pt Arial or Arial Narrow for chart callouts and labels.

There are no strict rules about choosing fonts for your presentation, but here are some tips I hope you find valuable:

Use a proportionally spaced font.

For attractive letterspacing and good readability, pick a *proportional* font—one that varies the space between the letters—for more attractive composition. There's just no good reason to use a *monospace* font—one with the same amount of space for each letter—for presentation-quality text. Monospace fonts can simplify alignment of numeric digits in the columns of a table. However, for precisely this reason, the numeric digits in proportional fonts are usually monospace, not proportional.

Proportional (Times New Roman)

`Monospace (Courier)`

Use sans serif fonts for chart labels, serif fonts for titles and explanatory text.

Plain, or *sans serif*, fonts like Arial and Helvetica are standard for chart text, such as axis and data labels.

The serifs, or curlicues and doodads, at the ends of letters can enhance the legibility of text, as well as add a distinctive impression. You might use a serif font, such as Times or Times New Roman, for chart titles and footnotes.

Don't intermix fonts within a chart.

You might use a serif font for the title, subheadings, and footnotes, and just one sans serif font for all the chart labels. Intermixing fonts within the chart itself will probably be a needless distraction for your audience.

Use upper- and lowercase letters.

While readability refers to the ease of recognizing words and sentences, legibility refers the ease of recognizing individual characters, or *letterforms.* Legibility is increased if you use upper- and lowercase letters. In particular, the *ascenders* and *descenders* of lowercase letters are an aid to quick recognition.

Ascenders	**Descenders**
b d f h k l t	**g j p q y**

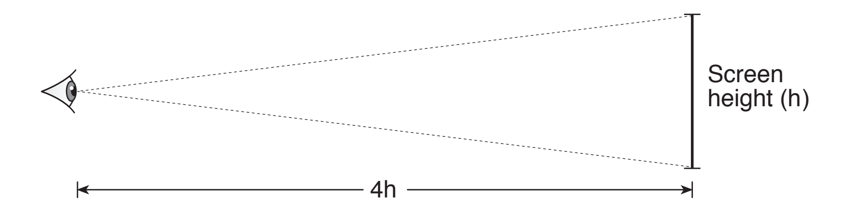
Liars Don't Care About Readability on the Screen

Whether your visual presentation will be viewed on a desktop monitor or on a wall screen, remember that the typical viewing distance is *four times the height of the screen.* Therefore, don't use text smaller than can be read comfortably at this distance.

Monospace Fonts: An Endangered Species

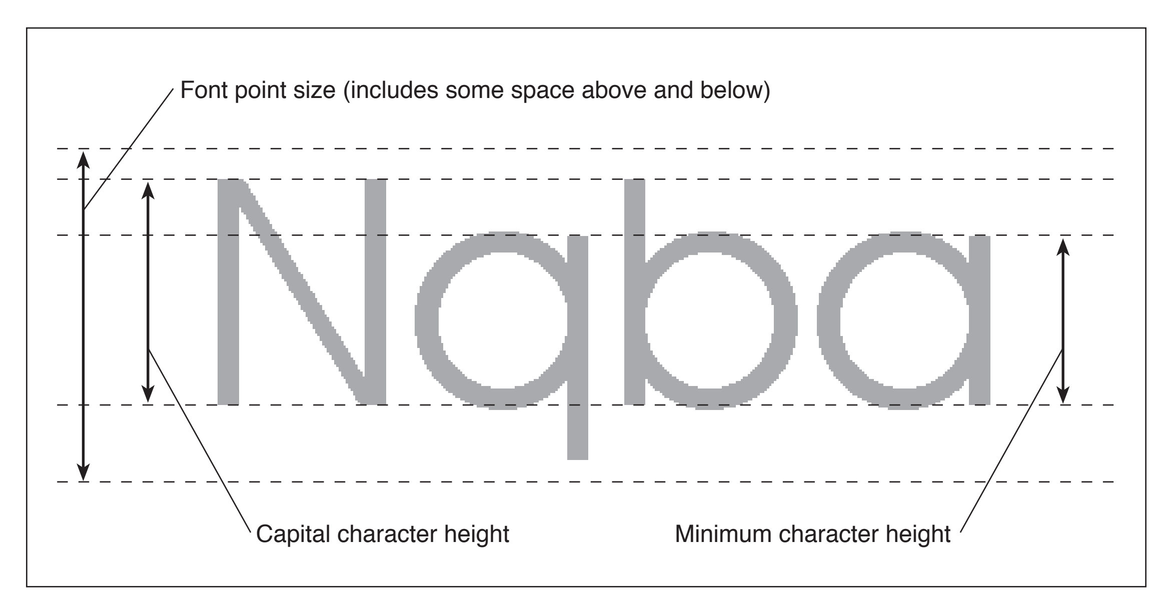
Are monospace fonts threatened with extinction? The only reason for the persistence of Courier and its clones in business correspondence is that people have entrenched expectations that go back at least as far as the 1960s and the first IBM Selectric typewriters. As far as I know, there are only two reasons to use monospace fonts—ever:

1. When you are building a table, monospacing makes it easy to align the entries vertically. (See Chapter 9 on setting a nice table!)
2. If you are writing a report or magazine article and must be concerned with the word count, monospacing can help. The number of characters per page will be more consistent than with proportional type. However, since word processing programs can count the words for you, what's the point? (Some programs give only a count of characters in a document. To get the word count, divide the character count by 8, unless you're using every long word in your vocabulary to write a scholarly dissertation.)



According to the Society of Motion Picture and Television Engineers, the minimum height of letters used in on-screen text should be at least 2.5 percent of the height of the screen. The height of a 14-inch computer or video screen is 8.25 inches, or 594 points (72 points per inch). So, the characters should be at least 2.5 percent of 594, or 14.85, points tall. However, remember that this measurement is *not* the type point size. The point size of a font is measured from the bottom of a lowercase descender to the top of an ascender, as shown in Figure 12.5. The specification here is for the minimum height of any letter, such as a lowercase *a*. The actual point size of a font for which the lowercase *a* is about 14 points tall would be about 20–24 points, depending on the design of the font.

Figure 12.5 *Don't confuse the minimum height of characters with the point size of the font.*



To further promote readability on the screen, avoid large areas of white as backgrounds. In general, use dark backgrounds with light-colored text for darkened rooms, such as formal boardrooms. Adjust these colors for the typical lighting conditions in the room. The brighter room light, the lighter the background color must be. In a highly lit room, you might need to use a light-colored background with dark text.



See Chapter 14 for more about use of color in presentations.

Just Your Style (Short Version)

Another mark of an accomplished liar is a studied sloppiness about punctuation and grammar, especially if the result is an ambiguous meaning or one that is downright wrong!

Here are some of the generally accepted rules of style that you can use for checking your presentation—both its charts and its report narrative. Some business organizations—and most publications—have their

own manuals of style, which cover many of these topics. Even if your company has a style manual, you will have to adhere to the editorial policies of the magazine or journal in published articles.

Punctuation

These aren't all the rules—just the ones that get missed most often:

Apostrophe

Generally, use *'s* to indicate the possessive form of a noun or indefinite pronoun, or add just the *'* to a plural form that ends in *s*. However, if the word ends in an *s* or *z* sound but is singular, add *'s* if this last syllable will be pronounced, but just the *'* if the possessive part isn't pronounced:

The box's shape
For goodness' sake
Mills' report
Mr. Ness's presentation

Colon

You can use a **:** instead of *and* to join a compound sentence. However, the second clause should amplify or clarify the first. Style books differ on whether to capitalize the first word of the second clause:

This is certain: we will have to
penetrate new markets next year.

or

This is certain: We will have to
penetrate new markets next year.

There is more reason to capitalize the second clause if it is lengthy.

Comma

The *serial comma* is often a point of contention. Is this sentence punctuated correctly?

We sell hardware, software, and service.

Or do you hold that the *and* takes the place of that last comma in a series:

We sell hardware, software and service.

I'd say experts on style are split about 50-50 on this one. Personally, I'm in favor of the serial comma, especially when the *and*'s start to proliferate:

We sell hardware, software, software support and hardware maintenance, and custom programming.

Another use of commas is to set off dates and geographic locations when they appear in the middle of a sentence:

We are located in Boulder, Colorado, but our home office is in Sioux City, Iowa.

On June 6, 2049, the note will become due.

However, if the reference includes *only the month and year,* the commas are omitted:

The note will come due in June 2049.

Dash

If a long dash (called an *em* dash) is used to separate two clauses, the comma that would otherwise go there is omitted:

Our technical advisors cannot afford
to make mistakes—our customers
will be watching us closely.

Ellipsis

This one is tricky. Use three dots (…) to indicate words that have been omitted from a quotation if the omission is *within* a sentence. If the omission is at the end, you need *four* dots—three for the ellipsis and one for the period! (Use the font's ellipsis character rather than a string of periods.)

Hyphen

Use hyphens in compound words (*self-examination*) or to divide broken words at the ends of lines. (Remember that the hyphen is being dropped increasingly from compound words, even if it means joining two vowels: *reenter*.) Use an *en dash* (slightly longer) instead of a hyphen to indicate a range (0–60 mph). Also, don't use a hyphen for a minus sign (−).

Double-hyphen

If the word you break at the end of a line in a report narrative was hyphenated at that point already, you should use a double hyphen (⸗), but you won't find it in most fonts. (This rule is seldom followed anymore.)

Parentheses

If an entire sentence is enclosed within parentheses, capitalize the first word and put the period inside the closing parenthesis:

(Information in this report is
based on unaudited results.)

But, if the parenthetical sentence appears inside another sentence, omit the initial cap and the period, although a question mark or exclamation point is permitted:

This fact is obvious (do we ever question
it?), but it is a dangerous assumption.

If the parenthetical expression is a fragment and it comes at the end of the sentence, put the period outside of the closing parenthesis:

This much is obvious (or maybe not).

Period

Style experts differ in their requirements for periods after some abbreviations. In general, use periods if the abbreviation is not an accepted acronym, but don't expect that rule to apply in all cases.

made in the USA
Meet me at 10 A.M. EST.
the Hon. Thurgood Marshall

Question Mark

Here are two different ways of posing a series of questions:

I want to know which product will you
specify? what size will you want?
do you prefer white or green?

I want to know (1) which product will you specify, (2) what size will you want, (3) do you prefer white or green?

Quotation Marks

Practice differs in the U.S. and in the UK, but here's the U.S. rule: Periods and commas go *inside* the closing quotation mark—even if not part of the quoted material. Colons and semicolons go outside. Dashes, question marks, and exclamation points go inside if part of the quotation, outside if part of the sentence as a whole.

If a quotation contains another quotation, enclose the embedded quotation in single quotation marks (‘ ’). If the embedded quotation comes at the end of the sentence, put the period first, then the single quotation mark, then the double:

Her exact response was, “I said, ‘Don't even think about bringing that up at the meeting.’”

Semicolon

You can use a semicolon in place of *and* to join a compound sentence, omitting the comma. But use the semicolon in place of the comma and keep the *and* if either clause contains commas:

Great Britain, France, and the United States were allies in both world wars; and, in the midst of intense economic competition, those ties still hold.

The formal style books also say to use a semicolon to separate compound items in a series:

The color schemes are red, white, and blue; orange, green, and brown; and blue, blue-green, and yellow.

However, some writers, especially popular journalists and novelists, are doing away with using semicolons in place of commas (omitting some of the commas, as well), with the idea perhaps that the flow of the sentence can be improved:

Great Britain, France, and the United States were allies in both world wars, and in the midst of intense economic competition those ties still hold.

Use of Italics

Use italics for emphasis; to highlight letters, words, and phrases referred to as such; in foreign words and phrases; and for titles of longer works such as books, manuals, and names of periodicals.

Use double quotation marks rather than italics for titles of articles within periodicals, as well as for chapters or sections within books.

Experts differ on which words are considered foreign and should therefore be italicized. A good working rule is that if you can find it in the A–Z part of the English dictionary, it need not be italicized. Examples are the Latin words versus and de facto, which are in everyday speech. But you might italicize a term if you think your readers won't know it:

She spoke of the *zeitgeist* of people at a dynamic company during a period of rapid change.

A common area of confusion is whether to put special words or phrases in italics or to enclose them in quotation marks. In the cases just cited, italics are correct. Use quotation marks (sparingly) to enclose words that are used in unusual ways. Avoid using quotation marks to give special emphasis to words that are used conventionally.

Capitalization

There are lots of capitalization pitfalls. Here are just a few of them:

Quoted Material

Capitalize the first word in a quotation *unless* the quotation completes the sentence:

The defendant said, "I made the whole thing up."

The defendant insisted that he "made the whole thing up."

Bulleted Items

The first word of bulleted items should be capitalized, whether or not the item is a sentence.

Headlines and Titles

This is an important one for publications and a point of some confusion.

Capitalize each word in a headline or title except for internal conjunctions (*and,* unless it's the first word), prepositions (*in, with*), and articles (*a, the*).

But—depending on the style expert—longer prepositions are also capitalized:

Choosing Between These Two Alternatives

One rule is, if the word has more than five letters, capitalize it.

Titles of Persons

Government and corporate titles are capitalized only if used with the proper name of the person:

President T. J. Mackenzie spoke.

The president spoke.

Some government officials can be addressed by their title in place of their names:

Mr. President, take a look at the polls!

Thank you, Ms. Speaker.

Numbers and Fractions

In a narrative and in chart notes, spell out numbers from *one* to *nine* unless they are used with units of measure:

We have two, perhaps three, choices.

The machined part will require
a taper of 3 cm.

Spell out the commonly used fractions *one-quarter* and *one-half,* unless the fraction is preceded by a whole-number digit or followed by a unit of measure. Omit the endings *st, rd, nd, th,* or *ths* after numeric digits or fractions. Do not use *of a* or *of an* between a fraction and its unit of measure (½ inch).

Use the typeset fractions included in a font rather than constructing them from two digits separated by a slash (½ rather than 1/2). And don’t put a space between a whole number and the fraction that follows it (5¼).

Alternate Spellings

Some people have specific preferences about alternate spellings, hoping perhaps to revolutionize the language single-handedly. For example, both of these spellings are correct, but a company style guide or publication should enforce the use of only one of them:

judgment

judgement

Where there are two correct spellings of a word, the preferred spelling is listed first in the dictionary. (But, which dictionary? And preferred by whom?)

Less obvious choices involve technical terms and jargon that are just starting to be accepted in general usage. Here are two versions of a familiar term that are now considered equally correct:

file name

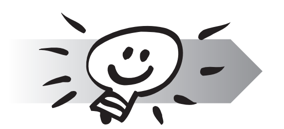
filename

$$$$INVESTOR'S TIP$$$$

Check it all for style and consistency with the financial proofreading guidelines presented in Chapter 13.

Until recently, you had other options for *workstation*—work-station and work station. Both are now considered incorrect.

A manual of style used by a corporation or publication will usually include a list of preferred spellings—especially for the technical terms and jargon of the trade or profession—and will also cite a specific edition of a dictionary to be consulted for anything else. The most commonly cited authority in the United States is *Merriam-Webster's Collegiate Dictionary.*



If you have to go to an unabridged dictionary to find a word, many of your readers might not know it. Consider using a more common synonym or even several simpler words that describe the concept.

In general, abbreviations should be avoided within sentences, but exceptions might be permitted for units (MHz for megahertz),

geographical names (Calif.), forms of address (Ms. with or Ms without the period), and so on. Style manuals should list the preferred forms of abbreviations and acronyms that are commonly used in the organization or publication.

Person and Voice

This category has less to do with charting but is nevertheless an important section of any corporate policy on report writing.

Depending on the formality of the report or presentation, a style manual might have rules about direct references to the author or to the reader. Some publications permit *I* and *you*, although many discourage *we* as condescending or confusing (who are *we,* anyway?). The "editorial *we*," with which the author attempts to hide in an anonymous group (with the corporate staff? with fellow wizards?), has all but disappeared from contemporary usage.



Liars are prone to the inclusive *we* as a way of including themselves with known experts or reliable authorities. Corporations are beginning to frown on the use of *we* in formal publications. Lest they appear condescending, presumptuous, or even adversarial, internal corporate publications must be particularly careful about this. It is often a matter of strict policy that *we* cannot be used unless it is clear from the context that the reader is included!

Some more formal publications recommend avoiding personal pronouns except for the third person—*he, she, it,* and *they*. However, this rule can force the author into stuffy formalisms such as

The reader will discover that...

A poor alternative is the cold impersonality—and vagueness—of the passive voice:

It will be discovered that...

Using the passive voice is all the more undesirable because most style manuals encourage writers to use the active voice wherever possible:

You will discover that...

Reading Level

Also related to style is the educational level that is assumed of the audience. Major-market newspapers assume a tenth grade reading level—which is also the standard for most undergraduate college textbooks!

Some grammar-checking software will analyze word-processing files and report the reading level as a score called the *readability index*. But here are some general guidelines, for which you don't need a grammar-checker:

- Avoid words having more than three syllables.
- Keep sentences to less than 20 words.
- Don't include more than two clauses in most sentences.

Want to Be Even More Stylish?

If you want to get serious about this style business, here are some of the most frequently consulted sources:

- Bernstein, Theodore M., *The Careful Writer: A Modern Guide to English Usage*
- Editorial Staff of the University of Chicago Press, *The Chicago Manual of Style*
- Fowler, Henry W., *A Dictionary of Modern English Usage*
- "Handbook of Style" in *Merriam-Webster's Collegiate Dictionary*
- Strunk, William Jr. and E.B. White, *Elements of Style*
- Warren, Thomas L., *Words into Type*

Other Quality-Control Steps

In addition to enforcing these standards, the writer of a report or the maker of a chart should review the material for *sense*, as in,

"Will this make sense to most of my audience?"

Although the content of your presentation may be reviewed in advance by your management, you should be responsible for assuring its accuracy. An important step is *fact-checking*—looking up statistics in reference books or contacting people who are quoted in an article or report to confirm the accuracy of the quotation.

Business Presentations Aren't Movies!

Oral presentations have a special set of rules. Whether you speak extemporaneously or give a prerecorded pitch, you must explain and interpret a sequence of pictures for the audience. You will be asking for trouble if you throw a chart of last quarter's sales results on the screen and then pause in silent meditation, attending the wisdom of the viewers. You must tell your audience what conclusions *you* draw from the data.

As the presenter, remember that the agenda should be yours. If your objectives aren't clear, or your conclusions aren't specific, you will lose control of the meeting, as well as your influence over the management decision you seek.

A time-honored technique for visualizing the structure of a presentation is to put the outline of your speech on the screen.

Sounds dull, you say?

I don't mean the whole outline, and not all at once. Rather, you can break up the points of a topic outline into text screens, which can give structure to a presentation. Show main topics as titles, and subtopics as lists.

Don't Sit at This Table!

A third type of text screen—the dreary table—is to be avoided. If you are tempted to show data in a table (or worse, if you try to put a whole spreadsheet on the screen), consider converting the numbers to a graph, which will almost always show the relationships better. For tips on creating charts—as well as tables, if you must—see Chapter 9.

Titles Are Topics

A title screen shows the text of a main topic, or title, sometimes with a subtopic, or subtitle, as shown in Figure 12.6. Insert a title screen before each section of charts in your presentation, just as you would introduce the paragraphs of a report with headings and subheadings.

Speechwriters call title screens *signposts*, for the obvious reason that a title points the way to what's ahead. The purpose is to reinforce the structure of your presentation so that the audience is with you at every step.

How I Spent the Summer

by Vincent Van Gogh

Impasto Enterprises, Inc.

Figure 12.6 *A title screen tells the audience what to expect, and follows the structure of main topics in your presentation outline.*

Put Your Agenda in a List

A set of related subtopics can be presented as a list, with the items bulleted or numbered, as shown in Figure 12.7. A list should recap the subordinate points of one main topic in your presentation outline.

However, showing a naked list (all items on the screen at once) encourages the so-called "read-ahead" problem. The audience can read faster than you can talk. They can see the whole list at a glance and will have a natural tendency not only to read ahead, but also to think ahead—perhaps anticipating your conclusions, perhaps wrongly!

The solution to the read-ahead problem is to subdivide a list into several separate screens as a *build-up sequence*—as shown in Figure 12.8. In choosing colors for the text, you can subdue, or dim, the previously presented items and highlight the current item. The subdued color might be a lighter shade of the background, and the highlight color should contrast sharply with it. So, if the background were dark blue, the subdued color might be light blue, and the highlight color might be white or yellow.

Figure 12.7 *A list recaps the details of one main topic in your speech outline.*

Figure 12.8 *A build-up sequence presents a list in stages to avoid the "read-ahead" problem.*

Subjects of My Paintings
In Southern France
• Landscapes
• Still life

Subjects of My Paintings
In Southern France
• Landscapes
• Still life
• Live models



A liar might not bother to signpost a presentation. After all, if the audience gets lost—so much the better! Or, how about including signposts for topics that are never discussed? (I guess I just slept through that part about the money-back guarantee.)

Take Your Cue from Your Text

If you're going to put words on the screen, you don't necessarily have to say them. But, if you do, use the language on the screen. Don't paraphrase. Otherwise, the result will be an annoying disconnect. And if you're going to say those words, say them as soon as the audience *sees* them. This synchronization of video with audio is called *visual cueing.* A picture should be cued, or should appear on the screen, at the moment that you begin to talk about it.

Careful speechwriters even make sure that the syntax, or word order, of the spoken words matches the text on the screen. For example, the audience might be distracted needlessly if they see

INCREASE MARKET AWARENESS

but hear you say,

"Recognition of our product in
the marketplace should be enhanced
in the coming months."

For the most effective reinforcement of your message, you should say,

"We want to **increase market awareness**
of our product in the months ahead."

Find a way to visualize all the important topics of your presentation and put them on the screen. Show key relationships in dramatic charts. But remember—those charts won't be persuasive without your honest and sincere words to back them up!

13 Financial Proofreading

Consistency Is More Than Its Own Reward

A **major pitfall of business reporting** is not only getting things wrong but also being inconsistent. Mechanical correctness is a must, and it's largely a matter of editing, of checking and cross-checking to get things right—and also to enforce consistency.

What's So Special About *Financial* Proofreading?

These types of documents, which may be a matter of routine for you, often involve presentation or discussion of business results:

- Periodic reports
- Management reports and analyses
- Letters to customers
- Narrative reports and whitepapers
- Presentations
- Email correspondence

Assuring the accuracy and reliability of this information goes beyond checking for mechanical correctness. And even at that level, seemingly small errors can be magnified in the minds of your readers because any kind of error invites criticism about the reliability of the rest of the document.

I haven't found a lot of resources on this topic. So I've included some of the issues I've encountered here, because creating financial documents has some unique challenges. And, like the issues that can arise when you're generating charts, ignoring these challenges can invite misunderstanding or give the wrong impression or even imply wrong information.

The overall impression your business report or presentation makes is a larger topic than just proofreading. It's not only about correctness—it's all about style. Style in a general sense is the kind of clothes you wear, the way you furnish your home, the way you set a table, the arrangements you make for your wedding. But in English composition, *style* has a narrower meaning. And this particular meaning of style encompasses everything—about which there are *no firm rules!*

Now, there *are* style rules. But this can be a contentious topic because those rules don't always apply, and not everyone applies them the same way. Let's just say, "Style is not black and white."

Large as the topic of style might be, achieving it comes down to enforcing consistency—which brings us right back to the tasks of editing and proofreading.

Style Guides

One big set of rules is found in *The Chicago Manual of Style.* This book started out as a guidebook that the University of Chicago made for its in-house press to deal with scholars and their publications. It rapidly became the bible for the book publishing industry. If you don't find the answer in there, nobody ever asked the question.

But we all know what happens when experts start making rules—they disagree. The other half of the publishing world—namely, newspapers, magazines, ad agencies, and public relations firms—use *The Associated Press Stylebook,* or AP Manual, which does not agree with the Chicago Manual on many obvious points, things that show up in your everyday writing.

The Serial Comma: An Important Point with a Twist

One everyday difference between the Chicago and AP manuals is the *serial comma* rule. Chicago wants a comma before the *and* in a series of items, and AP doesn't. Depending on which side of the Atlantic you hail from, this Chicago rule is also called the *Oxford comma* or the *Harvard comma.* Here's an example of consistency winning over policy: When veteran journalist Bob Woodward published his book *Obama's Wars* with Simon & Schuster, he omitted all the serial commas, even though the traditional rule in book publishing would be to include them.

So which rule book should you follow? I think it's amusing that on the cover of the AP Manual it says:

> *For journalists, students, editors, and writers in all professions.*

Not! Yes, reporters, publicists, ad agency copywriters follow this to the letter. But most book publishers don't.

There's a discussion "Just Your Style (Short Version)" in Chapter 12. If those tips are sufficient for your needs, perhaps you don't need to delve deeper. Suffice it to say, whatever you decide to do, strive to be consistent, and, at the very least, stick to your choices within the same document. (And, as I suggest later in this chapter, keep track of your choices in your own editorial style sheet.)

But if you have to comply with corporate standards or you're preparing documents for the public, you shouldn't stop there. And an extra level of care is justified where financial reports are concerned.

Perhaps you were taught that there's only one way, one set of rules, and you have to follow those. If so, your teacher had a bias, however he or she was taught. But teachers who are covering basic composition often don't point out these options. Perhaps they don't want to confuse their students. Or perhaps they themselves weren't taught that style isn't black and white.

There are actually many different sets of rules on style. To name just two organizations that differ, the American Medical Association (AMA) and the Institute of Electrical and Electronics Engineers (IEEE) publish their own guidebooks for editing scholarly papers in their fields. And they

disagree on many points. Most companies, and all large corporations, have their own in-house style guides. And it's not uncommon to find different sets of rules within the same organization, especially if it's a large enterprise that has many divisions or diverse product lines or is multinational. For example, it's not uncommon to find differing style guides for authoring aircraft repair manuals, annual reports, and advertising copy—all within the same company. Ideally, the company and its advertising and public relations agencies will decide on styles, and the agencies will help enforce them.

You probably know that the British have their own style rules, just as they insist on different spellings. And you can bet that organizations within the UK don't necessarily agree on application of those styles. (A Canadian author once remarked to me that British readers will tolerate American style rules, but Americans are put off by British prose. So in his books he uses American spellings and style rules and sincerely hopes his British audience will forgive him.)

In general, in U.S. English today, in the business world as in book publishing, the Chicago Manual is pretty much the standard. However, if you're working in journalism, including writing for the Web, the AP is the recognized authority.

Americans Don't Like that British Flavour (Two Examples)

American style:

This, that, or "the other."

British style:

This, that or 'the other'.

American style (the organization is a singular entity):

WidgetCorp today announces that it…

British style (the organization is a plural entity):

WidgetCorp today announce that they…

Differing Dictionaries

There's a dictionary in every office, although it may be buried in the back of the supply cabinet. Chances are, it's an old version and already out of date. I offer some examples of these changes in Chapter 12: *Workstation* used to be two words. Now it's uncommon to see it as anything but one word. As to *filename,* opinions are still divided. In IT, filename has been one word for a very long time. In general-purpose writing, you will often see it as two words.

Here are some newer entries, culled from the bleeding edge of linguistic evolution: *Podcast* may not be in the edition of the dictionary that's on your desk, but everyone knows about it. What about a *mobisode*? It's a video episode designed for viewing on a mobile device, such as a smartphone.

Other emerging practices will affect your style choices: If you were to write *email* with or without the hyphen, you are not literally wrong these days. *E-mail* is considered more formal, and that's the way you'll see it in some dictionaries. Lots of people, including many respected publications, are using it without the hyphen. I would venture to say within a few years it will be generally accepted *without* the hyphen. Many other words commonly found in your business correspondence and reports—especially technical terms and jargon—are similarly evolving.

In the back of many dictionaries, you will find an abridged style guide, which covers some of the same basic rules from the Chicago and AP manuals. And because these guides are written for general audiences, they typically try to have it both ways. Rules phrased like, "Some authors prefer…" are an acknowledgment that style is not hard and fast. There isn't just one way. After all, it's style. Is there a single dress code for work, no matter which company? Is there only one way to dress for a party?

Striving for Professional Appearance

Again, why bother with this stuff? It seems so trivial.

Whenever you make an appearance, and whenever you produce something in writing, you want it to be perceived as professional. And why not also—*classy?*

What are the elements of achieving that good impression? At a bare minimum, there must be no errors in content—no transposed digits, no misspellings, no wrong section or page numbers. That is, what you had in your source document agrees with what you show in your production document. (And if you find errors in the source document, check back with its originator before correcting them.)

But beyond having no transcription errors—and this is the real key to correctness—I repeat:

It's all about making choices and sticking to them—being consistent.

Why Is Consistency So Important?

Here is the formula I would call a dangerous equation. I don't mean to alarm, but the consequences are not pretty. The scary thing is:

INCONSISTENCY
= PERCEIVED SLOPPINESS
= LACK OF THOROUGHNESS

Any perceived lack of care begs the question, "How thoroughly did they check this report?" Or, "If I can't trust the smaller things like the way it's capitalized or punctuated, or the way my operating unit is abbreviated, can I really believe the numbers they're giving me?"

Ergo, the result of these perceptions:

= MISTRUST OF YOUR RESULTS!

That's a huge punishment for sinning about literally tiny stuff, maybe only a few stray commas. But numbers have commas—and decimal points. Doesn't it matter where you put those?

Enforcing Consistency

So how can you strive to be consistent? What are the kinds of things to look for when you're going over these documents one last time before they are duplicated or presented or delivered?

Here's a big one: Financial reporting periods should be consistent. If you grab financial data from corporate websites, from different customers, say, or from different reports, those sources could easily have *different fiscal years.* Year-end results probably will have been audited. But portions of years not yet closed will be unaudited. So, right away, if you're consolidating multiple statements, you could be mixing fiscal and calendar year-end figures, or audited and unaudited results. Such discrepancies should be disclosed (FYE = Fiscal Year End):

> Footnote:
>
> 3. All results are FYE (audited) except Tangible Net Worth, which is from the most recent three-month unaudited period.

Discrepancies can show up in currencies and exchange rates or units of measure: dollars and euros, ounces and grams, oil and water. Those pairs don't mix.

Other choices you should make and then stick to involve your manner of presentation. If you show a table to report one customer's results, you shouldn't show another customer's results on the next page in a bar graph. Perhaps you were in a hurry to meet a report deadline and pasted them from their sources because you didn't have time to generate new charts. But doing so is not only inconsistent, it's confusing. You should keep to the same presentation format unless you have a very special reason for highlighting one set of results a different way.

Consistency in document formatting extends to graphic design choices. Standardize on a font or a limited set of fonts. A good rule is to stick with one font for *body text*, also called *normal text*, and then use another font for headings.

Other potential inconsistencies can be in the legends and color codes of a document's charts. Make sure that legends use consistent names, spellings, and abbreviations for business units. Where possible, use the same set of color codes in all the charts. These types of errors will almost always creep in if you simply paste charts from multiple documents into your document.

Your company or your department should have standards for appearance of documents. It really gives a poor impression when sections authored by different workgroups use different fonts, margins, logo placement, and so on.

Being Careful with Abbreviations, Acronyms, and Company Names

If you have to abbreviate a word or a name, not only should it be an acceptable form of abbreviation, but also it should be abbreviated the same way every time you see it in the presentation.

Corporate parent-child relationships can be a source of confusion. You might start out talking about Sears, then later refer to Kmart. The two entities are parts of the same company now (as of this writing), but are you talking about separate brands or the same company? Kmart acquired Sears, but the consolidated company is called Sears Holdings Corporation. That's potentially confusing, and you shouldn't use the names interchangeably.

In another example, a credit report might consolidate data from a number of different agencies or databases. Those sources might not necessarily be consistent in the names they use to refer to the companies—especially if the reporting period spans the time of a merger or an acquisition.

Check the corporation's website to see how it displays its legal and brand names. The best source is to click the Terms and Conditions link, if there is one, at the bottom of the home page. And certainly if the readers of your report are your customers or shareholders of that company, they will be very sensitive to the correct use of the name.

Those Distinctive Company and Brand Names

Notice how these company names include unusual usage of capitalization and punctuation:

- *Wal-Mart* (later changed by the company to *Walmart*)
- *Kmart*
- *Toys "R" Us*
- *Linens 'N Things*

Proofreading Checklist

This section can be a basis for building your own editorial style sheet. Here are some areas of attention when proofreading and correcting your documents:

- Numeric consistency
- Punctuation and style of English composition
- Appearance, including graphic styles
- Accuracy

Numeric Consistency

Footnotes

Sometimes you won't be able to avoid inconsistencies in narratives or charts that consolidate data from multiple sources. Perhaps you can't reconcile audited and unaudited results. And in some cases it may not be practical or even advisable to convert all monetary amounts to dollars, since the rate of exchange might have fluctuated significantly since the figures were generated. Or your source data may be inconsistent in some other irreconcilable way. But you do have the mechanism of footnotes to tell the reader what is an exception and why.

Your management will probably set guidelines on which accounting periods to include in a particular report. In general, use a footnote to highlight any exceptional item in series of items, such as an unexpected or unusual result in a list of routine report items:

Note: These are mostly apples but here's an orange.

Use footnotes to clarify and qualify sources of data, estimates or budget figures instead of actuals, unusual circumstances, qualified auditor's opinions, or different accounting periods. In short, highlight any intentional or remarkable inconsistencies in the data.

In financial reporting, if the company has a relevant footnote in its financial statements, it may be advisable to pick up that footnote verbatim. It's usually wise not to rephrase it for legal reasons, so one of your proofing tasks would be going back to the source document to make sure the wording and punctuation are the same.

Truncation

If you have a column heading omitting thousands or millions, you want to make sure that is consistently marked, for example:

Net Sales

(000s omitted)

$12,345 = $12,345,000

In text, for round figures, you can use K for thousands:

$1K = $1,000

You can use *M* for millions, *B* for billions. (It's odd that the computer people use *G* for billions and the bankers use *B*, but that's just another stylistic divide that's not likely to be reconciled any time soon.)

Thousands separator

Always use a comma to separate thousands in monetary amounts:

$1,485

This form is preferred over $1.485K for exact figures.

Small decimal quantities

For amounts less than $1K, use the decimal point and add the two zeroes for cents, just so the reader knows you didn't accidentally leave *K* or *M* off the amount:

$349.00

instead of

$349

Otherwise the eye will stop at the number in a column of larger figures and wonder why it looks odd. Similarly, in technical writing, insert 0 as a whole-number placeholder in quantities less than 1:

0.004 sec

Negative numbers

Use of parentheses to show negative numbers is often inconsistent. Use curved parentheses, not the less-than greater-than signs:

(1,000) = –1,000

and is preferred instead of

<1,000>

which looks too much like

< 1,000

(If you get a stray space in there, you don't want the symbol misinterpreted.)

Minus signs also can be misinterpreted if there is any blank space between the sign and the digits. In general, it's not a good idea to use red text for negative numbers because red shows up as black on grayscale and black-and-white printouts. You might decide to use the parentheses as well as red, but then you're not strictly correct. Using red and the negative symbol (parentheses or minus) is a kind of double negative, but you'll see it, and if done consistently it's not necessarily incorrect.

Punctuation and Style

Bulleted lists

The grammatical principle of *parallel construction* is, when you have a list of items, you want them all to follow the same form, each presented the same way.

Begin a sentence with a capital letter and end it with a period, even if it's in a bulleted list. If it's a fragment, don't put a period at the end just to make the list consistent. As a rule, edit the phrases so they are all similarly composed sentences or fragments. That is, all lines are sentences and have periods or none of them do.

There are several other ways to handle bulleted lists, such as omitting periods at the end of each line even though the items are sentences:

- Do this first
- Do this second
- Do this third

Some authors insert a period after the last item in the list. A formal style, which is becoming somewhat old-fashioned is:

- Do this first;
- Do this second; and
- Do this third.

One very common way to distinguish between paragraph headings or bullet items, is an imperative—to use the second-person verb form:

- Connect the dots.
- Draw the chart.
- See the "Big Picture."

In this case, the subject of the sentence is understood: (You) Connect the dots.

Or use the gerund, the *-ing* form:

- Connecting the dots
- Drawing the chart
- Seeing the "Big Picture."

Connecting, drawing, and *seeing* in this example are nouns. So you'd put periods after the sentences, the imperative forms, but not after the noun forms, the gerunds.

Spacing after period

In touch-typing class, they used to teach students to put two spaces after a period. That's not done anymore. The extra space not only looks odd in word processed documents, but it can also throw off letterspacing, particularly when you are using full justification, where margins are aligned both right and left.

Incomplete sentences

How can you tell whether to treat a group of words as a sentence or as a phrase? Yes, a sentence has a noun and a verb. But in this example, some of the words are understood:

> Year-end results shown.

That's actually a sentence, or should be treated as a sentence:

> Year-end results (are) shown (here).

Title case

If you decide to initial-cap all important words, that's called *title case,* the usual style for titles of books and articles:

> *The Chicago Manual of Style*
>
> *The Associated Press Stylebook*
>
> *AMA Manual of Style: A Guide for Authors and Editors*
>
> *IEEE Editorial Style Manual*

And if you use title-cased phrases in a bulleted list, you want to make sure that the same capitalization rule is applied to each phrase in the list:

- Mechanical Correctness
- Appearance and Readability
- Section Numbering and Pagination
- Indexing
- List of Acronyms

Semicolon and colon

You can use a semicolon to join two sentences, taking the place of *and*. You can also use it to separate major items in a compound list that contains both major and minor items:

> The color combinations are red, black, and yellow; tan, green, and brown; and pink, purple, and blue.

A good alternative that's also more readable can be to insert *as well as* instead of the semicolon:

> The color scheme uses primary colors red, green, and blue, as well as the mixed values ochre, magenta, and taupe.

A proper use of the semicolon would be in month-end reports to separate groupings of receivable subtotals:

> Correct: Past-due amounts: $35,456; 4,928; 103,453

Accidentally omitting the semicolon or using a comma instead will be really confusing to the reader.

> *Confusing:* Past due amounts: $35,456, 4,928, 103,453

A colon can be used to join two sentences where the second one redefines or clarifies the first. You will find people disagreeing about whether the first word of the second phrase should be capitalized or not:

> One thing is certain: We will make our goals this year.

Whatever you decide, do it consistently.

In the following example, the colon is optional. But notice that it alerts the reader that a long list follows, and you're helping the eye find the phrases or items inside the list:

> There are a lot of choices: red-orange, yellow-green, blue-green, yellow-orange, or gray-violet.

If it's a very long list, the correct form would be a number with a parenthesis after it before each list item. Doing it this way is a good alternative to a bulleted or numbered list when you don't have much vertical space:

> There are a lot of choices: 1) red-orange evoking sunsets, 2) yellow-green with a spring feeling, 3) blue-green for a deep-sea experience, 4) yellow-orange sunburst, or 6) the gray-violet of evening.

Use of the hyphen

Dictionaries and style books often disagree about the treatment of certain compound words. For example, some financial wizards spell the noun or adjective form *writeoff* as one word, and some hyphenate it *write-off*. But no matter which of those forms you use, if it's used as a verb, the correct form is two words:

> We will write off the batch in July.

Other terms that have similar variations are *setup, followup,* and *breakout.*

Apostrophe

Here's a very common mistake with the apostrophe. The word *its* is an exception to the rule for putting an apostrophe to show a possessive. *It's* with the apostrophe always means *it is.*

A similar error, when you pluralize an acronym or a number or year shown in digits, such as thousands, you aren't supposed to use an apostrophe:

> ABCs
>
> 000s
>
> The 1900s

You just add a lowercase *s* to indicate the plural. (Chicago and AP agree on this one!) However, when you do the same with letters, the apostrophe is necessary to help readability:

Mind your p's and q's.

Smart quotes

This is more of a Microsoft Word thing: Those smart, or curly, quotation marks can drive you nuts, especially if you're transferring documents between platforms such as PC and Mac or Word and OpenOffice. But also if you just insert an apostrophe to indicate a contraction, such as *'90s,* Word is going to think it's the start of a phrase, and the curl will go the wrong way. To avoid the problem, you can use the Insert > Symbol command and select the correct character.

Automated checking

Use the automated spelling and grammar checking features in the software as a quality-control step, but nothing substitutes for checking it all carefully by eye. If you're in a hurry and you don't want to consult the dictionary, keep a browser window open and do a search on the word, especially for proper, company, and brand names. Doing an online search isn't always reliable, so if you see variations in your search results, you'd better check a more authoritative source.

Abbreviations in sentences

Using abbreviations within sentences is always a bad idea. An exception would be labels in tables and charts, when you don't have room. In those cases, use standard abbreviations, if possible:

mgmt	management
add'l	additional
pgm	program
stmt	statement

In charts, tables, and figures, it is acceptable to omit periods after abbreviations. However, place a period after the abbreviation if the reader might not recognize it as such.

You can use acronyms in your sentences, but don't use periods with them and spell them out on first use:

> The advertising rates for Cost Per Click (CPC) have increased steadily over the last quarter. Our ability to leverage CPC in the future will depend on our success at converting click-thrus to purchases.

An exception to using or not using periods with abbreviations is to follow legal forms for company or brand names, even if the form is not consistent with your style rules. For example, some company names use a period after *Inc, Co,* or *Corp* and some don't. However, the abbreviations *LLC* or *LLP* never use periods.

Symbols

Also marked on first use would be trademarks—the TM symbol *Acme*™— and registered trademarks—R in a circle *Acme*®. (There is a legal distinction between the two. Check the brand owner's site for the correct one.)

Standard abbreviations

If you're going to use abbreviations in chart labels, it's cleaner and permissible to omit the periods. Just make sure you are consistent. An exception would be an abbreviation that people are not used to seeing—use the period to indicate it's a deliberate abbreviation and not a typo.

The ampersand (&)

Never, never, never use & in a sentence. The only exception is when it occurs in the legal form of a company name:

> XY&Co

or in a standardized acronym:

> P&L (for Profit and Loss)

This is a common mistake and always looks awful. And it only saves you two letterspaces over spelling it out. It's okay to use the ampersand and other symbols in chart labels and column headings, particularly to save space and especially when you're forced to use small font (*e.g.,* 8pt) to fit a large table onto the page.

Appearance

Paragraph and Character Styles

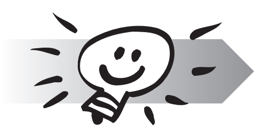
Don't intermix fonts randomly. The Normal template in Microsoft Word has basic styles such as:

Normal

Heading 1

Heading 2

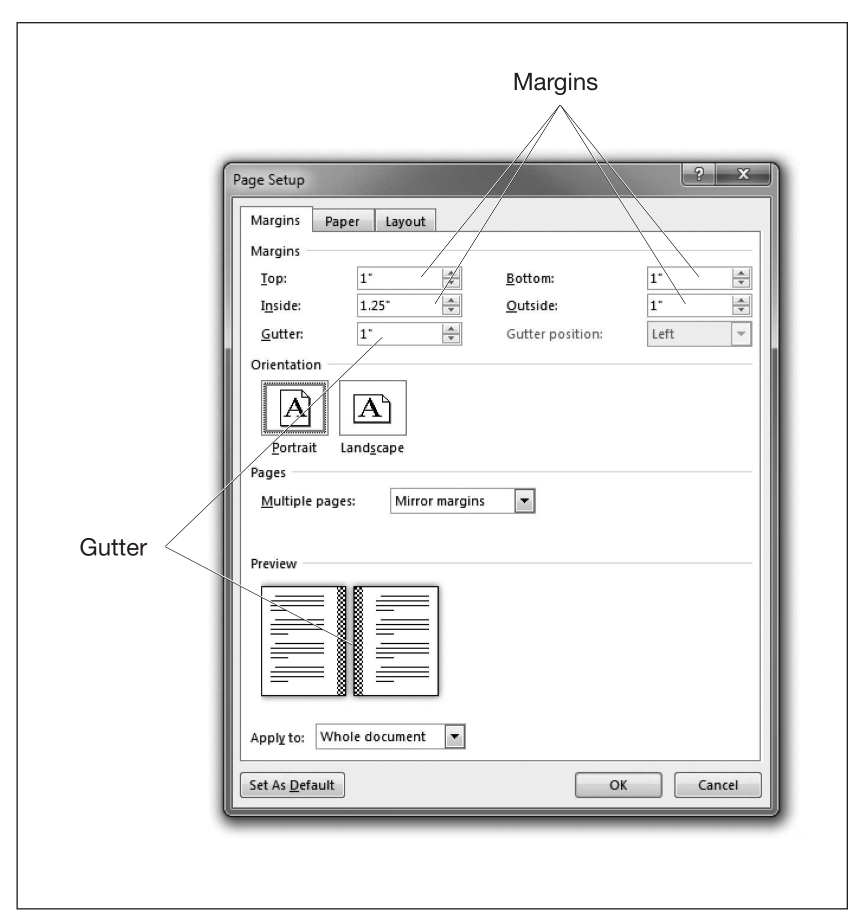
If you start creating your own styles manually, rather than applying readymade styles, the program will add them as new styles (unless you turn this feature off). The result can be a mish-mosh. You'll have more attractive and consistent documents if you learn how to apply styles and keep them from multiplying like rabbits. Documents that have multiple authors and editors almost always have this problem. Good practice for a final edit is to reformat the entire document, using only the styles that you wish to appear.



It's good practice not to use the Normal style in Word for default paragraph text. That's because most, if not all, other styles are typically variations on Normal. Create a new explicit style, such as Report Body Text that is not based on Normal. Doing so will make it less likely that any redefinition of the Normal style in other users' Normal.dotx templates will cause the appearance of your document to change when they open it.

There's a practical reason to use heading styles correctly. When you go to insert a table of contents in Word, the application is only going to pick up those headings that are actually formatted with one of the heading styles. That is, if the topic is formatted as Heading 1, it will be picked up as a first-level topic in the contents. But if it's simply been boldfaced manually and changed to larger point size, the phrase will be ignored.

Figure 13.1 *Microsoft Word provides these settings for controlling page margins and the* gutter, *or extra space along the spine for binding.*



Page margins

Also in terms of appearance, make sure your page margins are consistent. If the document is three-hole punch or has some type of binding, make sure you've added enough extra space, or gutter, on the bound edge (Figure 13.1).

Try to apply consistent page formatting even if you are assembling multiple documents from different sources. Resize pages to fit within your margins, unless you absolutely must preserve the appearance of the original document.

Page numbering

Don't omit page numbers, which are not only helpful to the reader but also necessary if you're going to insert a table of contents. Page numbering can be tricky in Word, especially if your document has different sections. But learning how to insert section and page numbering correctly, along with headers and footers, is a skill worth acquiring.

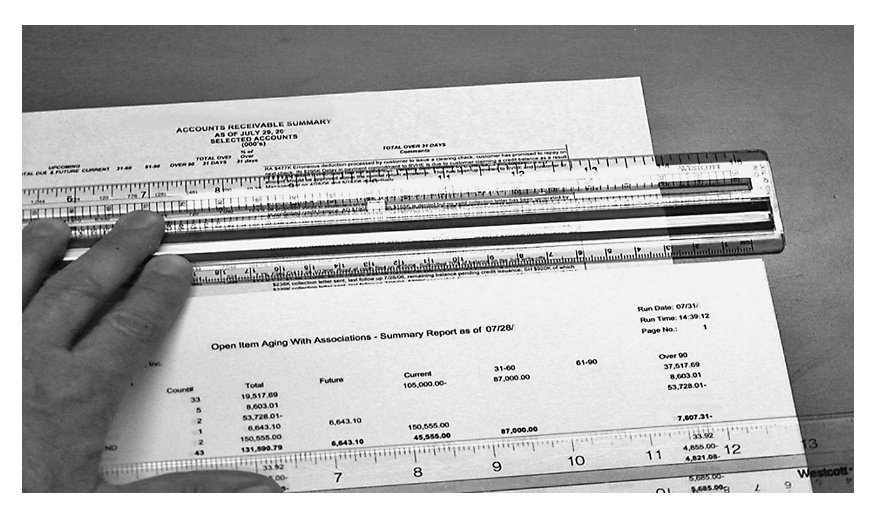
Accuracy

Proofreading technique

There are two basic ways to proof, one informal, the other formal. If you're proofing by yourself, the recommended technique is to look at hardcopies on your desktop. Use two rulers, one on the source document, the other on the printout of your report (Figure 13.2).

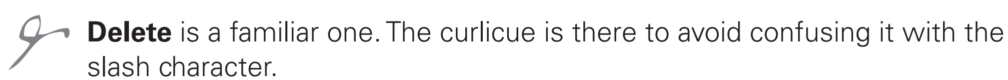
Proofing in the legal profession and in the publishing industry can be really very old fashioned. It takes two people, one to read the source document and the other to follow along in the production printout, making particular note of editor's or proofreader's corrections and markups, and spelling out proper names and unfamiliar words.

Figure 13.2 *Proofreading by yourself can be done with hardcopies and rulers to make line-by-line comparison. (Photo by the author.)*



Proofreader's Marks

You'll find a listing of proofreader's marks on page 100 of the 15th Edition of the Chicago Manual. (Do you think they did that on purpose?) Not everybody uses the same ones, but it's important if you're going to be working as a team to standardize on one set of marks.



Delete is a familiar one. The curlicue is there to avoid confusing it with the slash character.

For a **period**, put a circle around it so everyone knows it's not just a stray mark on the page.

Comma has a caret (tent-shaped symbol) on the top of it so it's distinguished from an **apostrophe**, which has the inverted tent underneath. And the mark to insert a comma is typically shown below the line of text, apostrophe above, just to be clear. The caret indicates where in a line the insertion should go.

The arc over the **hyphenation mark** indicates that the space between the two words should be closed up around the hyphen.

One slash through a character means it's to be made **lowercase**. If you're actually inserting a slash character, it's shown as a double slash.

A triple line underneath a character or word means to **capitalize** it.

A bracket indicates **indent** (or outdent) depending on its direction and placement adjacent to the line of type.

The **colon** symbol has carets above and below, to make it obviously distinct from the comma, the apostrophe, and the period.

The insert **semicolon** mark is markedly different from the colon symbol.

Accountants and editors use the symbol *s/b* to mean ***should be***. It's a good shorthand for editing because the thing that's wrong is always first, the correction always second. Engineers use *was* and *is* for the same reason.

A long dash is an ***em* dash** because it's the width of the letter *M*.

A short dash is an ***en* dash** because it's the width of the letter *N*. Now, it's not the *very short* dash between words, because that's a hyphen.

The **paragraph** symbol is very familiar.

The **transpose** symbol is used to change the order of letters, words, phrases, or sentences.

Editor's style sheet

Keep a style sheet for your work group or department, especially if you don't have a corporate style guide. Even if you're following the Chicago Manual, it can be a cumbersome book. It's handy to have a concise list of the styles and rules that you use often. And, as a practical matter, once you make a choice, you might otherwise forget it by the time you've finished editing a long document.

Tips for Email Correspondence

Email is an informal medium and not subject to all the rules, but here are some useful tips.

- Try to use complete sentences. You want to avoid misunderstandings. Remember, email is a legal record, and it can be subpoenaed as evidence in court. If you're using sentence fragments and your thought can be misconstrued, it probably will be.
- Avoid what I call *comic book writing*: including emoticons, symbols, and punctuation marks used in nonstandard ways. Don't overuse bold, italic, and underline. Even one exclamation point is usually one too many, especially in business documents. This form of "creative writing" only looks cute to the people who do it.
- And as everyone knows but the people who do it, expressing yourself in ALL CAPS is just considered rude.
- Document important names, dates, and events in your communications just as you would in a formal letter. Of course, the time stamp and the header information do some of this for you automatically. If a third party read your note, would she know what it's about? Granted, it's probably part of a message thread. But if you refer to "it," will your meaning be clear if the correspondence is reread a month from now? With just a few more words, you can set the record straight:

> Martha,
>
> It is nowhere near what we discussed. I have a real problem with it, and unless you can justify it, my supervisor certainly won't approve.

Suggested revision:

> Martha,
>
> ~~It~~ **The amount due** is nowhere near what we discussed. I have a real problem with ~~it~~ **the low discount percentage,** and unless you can justify ~~it~~ **your calculation,** my supervisor certainly won't approve **payment.**

- Use your auto-signature block if the correspondence is going to the outside. It provides your contact information, which is useful especially if the document is printed out. The block should also have disclaimer language if your firm requires it and is particularly appropriate if the communication contains information that could in any way be construed as confidential or proprietary.
- When using your mobile device, your messages are likely to be concise, terse, even cryptic. It's good practice, then, to confirm with a longer email or telephone conversation when you're back in the office. Start with, "As I mentioned in the very brief note I sent you from my mobile…" and then explain more fully.

Financial Proofreading Goals

To recap the advice in this chapter, checking your documents has two essential goals:

No Errors

Your basic goal in proofreading is to make sure that the source documents and the production document agree, especially numerically. Did we drop a digit? Are all amounts in the proper columns? Is there a

human error that caused the documents to mismatch? Do the formulas in spreadsheets always generate valid results?

Professional Impression

But beyond correctness, you want the document to have consistency, which conveys professionalism. And it's not that the presence of consistency is drop-dead impressive. It's that the lack of consistency is so obvious and so obviously unprofessional. You and your organization simply can't afford to look sloppy.

You want your reports to give the same impression as your best business dress:

PROFESSIONAL AND CLASSY!

14

Color

Its Uses and Abuses

$$$$INVESTOR'S TIP$$$$

Some purists discourage the use of color in financial reporting, favoring grayscale exclusively. That's somewhat extreme, but the fact that they're worried about the uses and abuses of color should tell you something.

As with so many other lessons in life, things that are pretty can be especially deceiving. Applying color to your charts might be aesthetically pleasing, but looking pretty shouldn't be your only thought. Color is a powerful tool for communication—and by now you should appreciate that any graphic tool holds the potential for abuse by shameless prevaricators. In fact, some accounting professionals are so fearful of the potential misuses of color that they advise using only black and white!

Your world needn't be so dull. Go ahead, splash on the color. Just be aware of the attentions you invite when you try to dazzle your admirers with it.

As in personal grooming or interior decorating, you need to develop a sense of style when using color in your charts. As with the other graphic tools I've discussed, learning to use color effectively involves both knowing the expectations of your audience and understanding the circumstances that surround your presentation. You need to know subjective things like people's biases about color, as well as some objective facts about how colors are reproduced in different media—such as paper, video, and Web page.

In the realm of color, there's a lot for the ambitious liar to exploit. And, as with subjective notions such as screen direction (left-right, up-down), our prejudices about color are so ingrained that most of us aren't aware of them.

You can also think about color as an aspect of product design—the product being the presentation you're preparing. Initially, your natural inclination will be to please yourself, to make selections that suit your personal taste. But there will almost always be other opinions besides your own that you can't ignore. Your priority should always be meeting the expectations of your audience.

The situation gets more complicated if you are preparing a presentation that will be delivered by someone else—your boss or perhaps a client. Here you might run into the touchy situation where your client's preferences about color run counter to his or her best interests—as far as getting a desired result from the audience is concerned.

A factor that can affect all these sets of opinions is corporate style. Most business organizations have adopted specific sets of corporate colors, possibly also including required typefaces and other graphic rules intended to create a unique, identifiable image. Often these rules are codified in style manuals, which must be followed in preparing any major presentation to management or the public. As a result, if different companies are involved, you might need to create separate sets of colors and presentation styles for each project or client.

Most people don't have to be so exacting about color. They needn't worry, for example, that referring to the *shade* of a color is not the same as describing its *tone.* But if you want to develop a vocabulary for using color in presentations, you must pick your words more carefully.

Color Terminology: The Short Course

The word *color* is too all-encompassing to describe the rich variety of visual impressions that surrounds us. Just as Greenlanders have many different words for *snow*, graphic artists have developed more precise terms for dealing with color. Here are some basic definitions. (The mixing of these elements to produce different colors is diagrammed in Figure 14.1.)

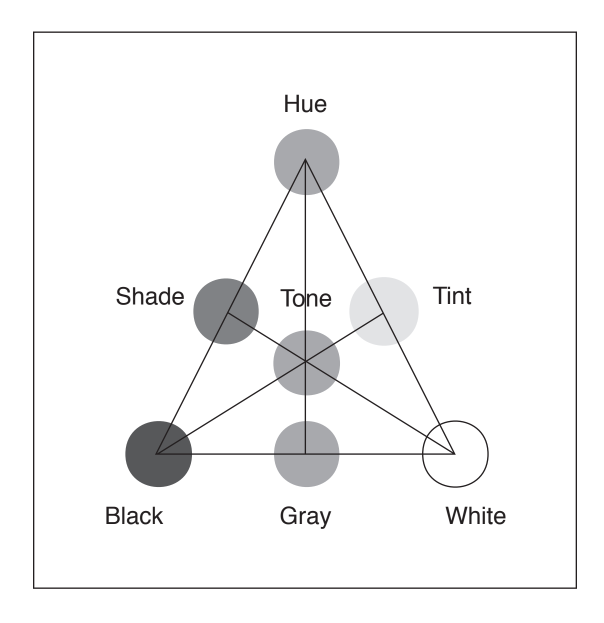


Figure 14.1 *To the discerning graphic artist, what the rest of us call a* color *is a* tone *that has at least six different contributing factors.*

Hue

Describes the primary colors from which a specific color blend, or tone, can be made. The basic hues are red, red-orange, orange, yellow-orange, yellow, yellow-green, green, blue-green, blue, blue-violet, violet, and red-violet.

Tint

Refers to the amount of white that is mixed with a color. Light tints appear washed out. A technical term for tint is *chroma*, or color purity. Some color experts prefer to call it *saturation*.

Shade

Is the amount of black, or darkness, in a color. Similar terms are *value, lightness,* and *intensity*.

Tone

Is what in everyday speech you'd call a *color*—the combination of hue, tint (saturation), and shade (value).

The three elements of computer-generated colors are Red, Green, and Blue, or *RGB*. But trying to get the color combination you want by adjusting R, G, and B controls isn't so easy. So, most computer graphics software translates RGB into a more intuitive set of controls: Hue, Saturation, and Lightness, or *HSL*. In some programs, these terms are called Hue, Chroma, and Value (HCV), or Hue, Saturation, and Value (HSV).

The Psychology of Color

Colors not only mean different things to different people, but the meanings can also be different depending on the situation those people are in!

At the most subjective, nonverbal level is the question of *mood,* or emotional response. Consider the variations in mood of some basic color hues for different audiences. A survey of American movie audiences reveals one set of moods, as shown in the second column of Table 14.1.

Similar tables are used by movie technicians to adjust the overall tint of scenes to better fit the mood the director intended. Now, you might think that such minor differences don't have anything to do with the bold colors in business presentations. But I'd say that the effect is analogous—in the case of the movies, the color difference is very subtle, and so is the effect on the audience. Make the colors bolder, and you will magnify their effect.

Now look at the third column of Table 14.1. There's a completely different set of preferences for financial managers, based on how they ask professional designers to create presentations for them.

The differences don't stop there. Doctors and other health care professionals, shown in the fourth column of the table, don't agree with the financial types.

Table 14.1 *Subjective Interpretation of Color by Different Audiences*

Color	Movie Audience	Financial Managers	Health Care Professionals	Control Engineers
	INTERPRETATION			
Blue	Tender	Corporate, reliable	Dead	Cold, water
Cyan	Leisurely	Cool, subdued	Cyanotic, deprived of oxygen	Steam
Green	Playful	Profitable	Infected, bilious	Nominal, safe
Yellow	Happy	Highlighted item, important	Jaundiced	Caution
Red	Exciting	Unprofitable	Healthy	Danger
Magenta	Sad	Wealthy	Cause for concern	Hot, radioactive

In the fifth column is another set of reactions from a group of control engineers with responsibility for monitoring a nuclear reactor.

It should be apparent from the table that some biases about color have to do with a person's professional and cultural background. But notions about color aren't that simple: Our interpretation of colors can change with the situation. People make these mental shifts frequently and without conscious effort, usually depending on their understanding of the *purpose* of a visual message.

For example, in some circumstances, it will be natural for movie audiences to react more like control engineers. When the moviegoers exit the theater parking lot in their cars and approach the first traffic signal, their situation will be much closer to that of the control engineers

in a power plant: Very quickly, they must interpret a limited set of color codes. In this simplified set of codes, a red light always means *stop*, regardless of how the person might "feel" about red in some other situation.

On the other hand, people with specialized biases, like the control engineers, do not necessarily carry them into other aspects of their lives: A control engineer's reactions to colors in the movie theater are not necessarily any different from those of doctors, accountants, or the public at large.

What does this mean to you as you design a business presentation? Again, you must tailor the presentation to the biases of your audience. Be aware of these biases. To an accountant, green money is good and red ink is bad. Doctors want it just the other way around: oxygen-rich red blood is very good, green necrosis very, very bad! But the situation is also important. The accountants and the doctors will agree on what to do when they see traffic lights on the drive home. And individual tastes and moods will vary even more, for all kinds of subjective reasons—such as the color of the sky on the day you make your pitch!



Liars are skilled at sensing the mood of an audience and then manipulating it to their own ends. A keen sense of how colors affect moods is therefore an essential skill. Yes, perhaps using corporate blue is usually a safe plan. But who knew that using a lot of red in making an investment pitch to a group of doctors would stimulate more excitement than alarm?

What Do You Mean by That Color?

Creating a mood with color involves making an overall impression, perhaps an unconscious one, on the audience. But beyond mood, which can arise even from unintentional color choices, there can be deliberate and obvious uses of color in presentations. That is, color can be used for specific communication purposes.

Color can be used for 1) a kind of coding, 2) rendering of real-life pictures and artistic expression, 3) as a design element of graphic materials, including charts. Business presentations often include all three types of uses.

Color as Coding

The situation of the control engineers in the nuclear plant is an example of color used as coding. Computer displays that monitor conditions within the reactor may use specific colors to signal the status of each of the plant's subsystems.

The power of any coding scheme is its simplicity. When colors are used as codes, you must *limit the choices to the fewest colors necessary*. The greater the number of colors, the more likely your audience will become confused.

For example, think about how confusing it would be to deal with traffic signals that displayed six or seven colors. Instead of GO … CAUTION … STOP, you might be confronted with color codes for:

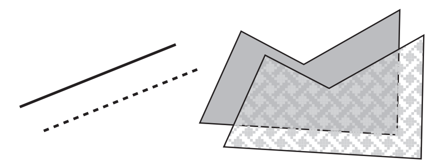
GO

NO LEFT TURN
NO RIGHT TURN
NO U TURN
CAUTION
WAIT FOR PEDESTRIANS

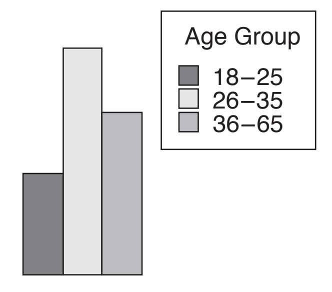
When colors are used as codes, the precise shades or tints of the color hues probably won't be important—as long as the audience can easily distinguish one code from another.

In developing codes, remember that some people are color blind. The common difficulty is distinguishing opposite colors on the color wheel (more about that later). Red and green might both look like gray or some other neutral color. A solution to this problem is to code the objects two ways—by pattern as well as by color. (This is sometimes done on traffic

lights. If you look closely, you might see a bar across the red lens.) So, even if you code two lines in a chart separately by color as red and green, you might also distinguish them as solid and dashed, and areas as solid and patterned:



An example of color used as coding is the legend of a chart. The legend provides a key that matches colors, patterns, or both to bars, lines, areas, or pie slices:





A liar might well use too-similar colors in the legend of a chart—attempting, for example, to blur the distinction between otherwise unlabeled plots of profit and loss.

Color as a Rendering Tool

The color tones in a photographic scene might number in the thousands—from a possible range in nature that might number in the millions. From the standpoint of the range of color choice, a fine artist would have the same requirement. You wouldn't want to restrict the color choices of a budding Picasso to a few simple codes. Use of color to reproduce a scene in nature or as artistic expression would require the widest possible range of colors (see Figure 14.2). Notice that simply reproducing the original picture as a monochrome photograph not only loses the richness of full color but also makes the scene lack depth.

Figure 14.2 *Rendering a scene from nature requires the widest possible selection of colors—the opposite of the requirement when color is used as coding. Admittedly, this gray-scale image is neither realistic nor pretty. But consider how much information it* lacks. *(Dogs, I'm told, won't see the point of this comment.)*

(Photo by Gary Palmatier.)

The need for richness and flexibility in color selection is just the opposite from the requirement for color codes. Instead of restricting the number of colors, you need to have all the possible subtleties of hue, shading, and tint.



The world of product photography for advertising is full of liar's tricks. To heighten its visual appeal, food is often photographed through a pink filter, which conveys the feeling of warmth. Nobody wants a hamburger that looks slightly blue or green! In some cases, selective retouching will be required—even though you might add green to make that salad to look cool and crisp, you don't want its surroundings to look chilly.

Color as a Design Element

When you are designing a business presentation, the range of colors you need will fall somewhere between the extremes of natural color and color as coding.

In general, your basic approach will be to use colors as codes, as in keying lines and bars to their respective company divisions or product lines. For these chart elements, the rule about coding applies: Keep the number of codes to a minimum to avoid confusing the audience, and make sure that the color tones you use are distinctly different.

Color can also be used as *thematic* coding—to distinguish different sections of a longer presentation. For example, the marketing presentation might have a blue background, followed by the financial section in green. Although the range of colors here might be considerably wider than the basic codes you've used in charts, you should still try to limit the number of color themes. Too many color themes—which are, after all, a type of coding—will confuse the audience. An effective technique is to build each major section of a presentation around a basic color hue, then vary the tints or shades of that hue within the section.

Your overall objective in choosing colors should be to reinforce and clarify the content of your presentation. You should think of color as one of many essential design elements such as font, line weight, and spatial composition. Used in coordination, these elements can produce a visual composition that is readable, understandable, and aesthetically pleasing.

Color Range in Business Presentations

Because too many color codes or color themes can confuse your audience, there is no obvious need for a wide range of color in the presentation itself. Nevertheless, you will need a wide range of colors from which to choose—enough variety to permit you to select a few precise tones that work well together. When your needs move beyond coding to include pleasing and effective color composition, the range of colors needed actually becomes quite large. To understand why this is so, you must understand some of the factors that determine color range.

"How many colors do I need?"

Even the most inexpensive devices these days are capable of displaying literally millions of colors. However, unless you're rendering live photography or videography—where a full range of subtle variation really counts—limiting your choices is actually a good idea. Have a look at the following section on color schemes to find out how many colors are enough.

Color Scheme, Palette, and Gamut

Your color choices for a given visual composition make up the *color scheme* of the composition. Within a computer or in a software application, the set of specific colors available is called a *color palette*. The range of colors that can be reproduced on a specific device, such as a video projector or a color printer, is called its *color gamut* (see Figure 14.3).

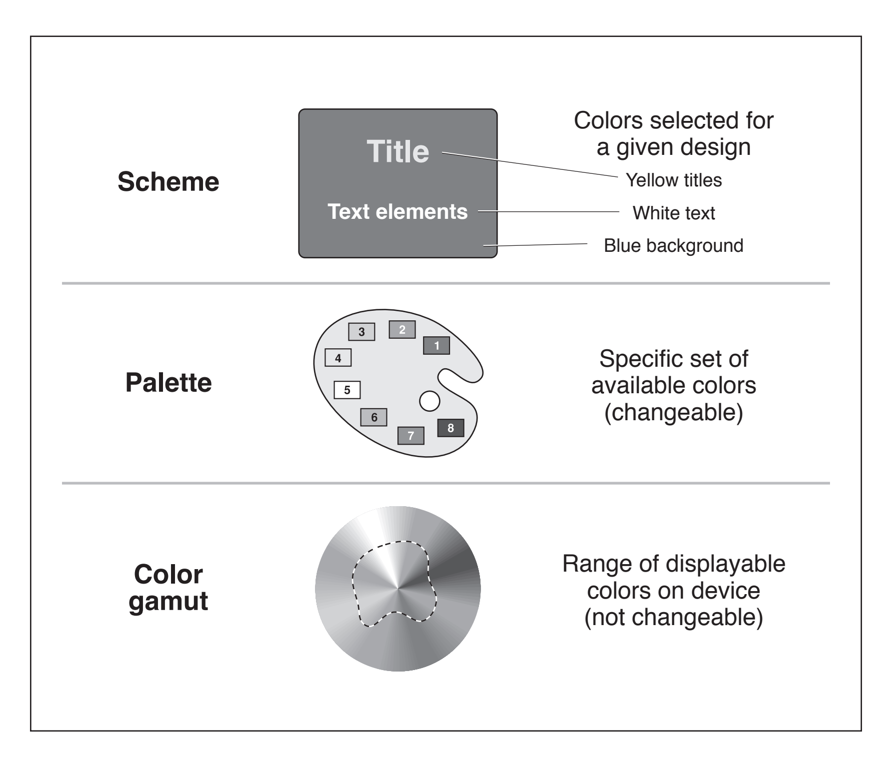


Figure 14.3 *Here are some more precise terms for color selection.*

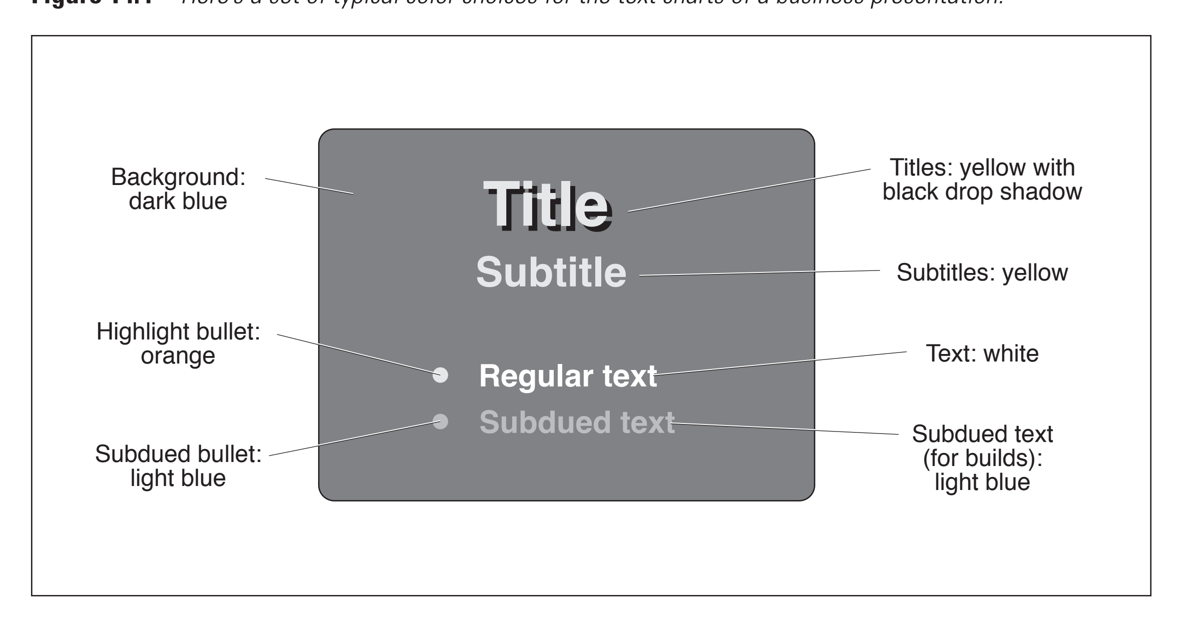
A color scheme is an abstract idea—your notion as a designer of how a presentation should look. A color palette is a tool for carrying out a color scheme. The color gamuts of computer devices are physical, technical characteristics that determine which color tones you can choose for your palette.

So, when you are choosing a palette for use with graphics software, you need to think about both your ideal color scheme as well as the color-gamut constraints of the computer. Your color selections, therefore, will be a compromise between your desired color scheme and the color gamut that is available to you.

Color Selection

As shown in Figure 14.4, your preliminary decisions about a color scheme for your charts should include selections for background, main title and subtitles, text (explanatory notes, lists, and chart labels), highlighted text, and subdued text. (For an example of build-and-subdue technique for text charts, see Figure 12.8.)

Figure 14.4 *Here's a set of typical color choices for the text charts of a business presentation.*





I can't emphasize it too strongly: *Background color will be your single most important choice.*

The background color you pick will largely determine the color scheme and the color palette you use to carry out that scheme. The specific tones in your color palette will be influenced by the type of output medium (such as paper or video screen), room lighting conditions, the mood you want to create, and whether your presentation will also be published in other media, such as on paper handouts.

Remember that the value of a color refers to its degree of darkness. The value of the background in a color scheme determines the *contrast range* of the entire picture—the range of values between the lightest and the darkest color tones. The greatest contrast generally should be between the background and the titles or text—because *contrast promotes readability*. Again, the controlling factors include the display medium, lighting conditions, mood, and reproduction requirements.

The most striking color graphics for formal presentations have relatively dark backgrounds. But, you should increase the value of the background shade in proportion to the brightness of the light in the meeting room. Brighter room lights require lighter background colors and a narrower contrast range, as between background and text.

So, the colors you choose for titles and text should be among the lightest colors. You should pick subdued, or dimmed, color shades for toning-down or deemphasizing certain lines of text or areas of a chart. An appropriate subdued color for slides would be a lighter shade of the background. Highlight color is used to draw attention to specific elements or objects. This color should be the brightest of the colors in the chosen color palette. Examples of the same chart designed for different media and display conditions are shown in Figure 14.5.

In brightly lit rooms, the opposite is true. The most effective images will have light-colored backgrounds. While formal presentations usually are done in darkened rooms, you might prefer to leave the room lights on, as in training sessions or demonstrations—or just when you want the audience to be able to see you clearly. The brighter the lights in the room, the brighter or lighter the color of the background must be. Titles and text are usually rendered in contrasting, but equally vivid, colors.

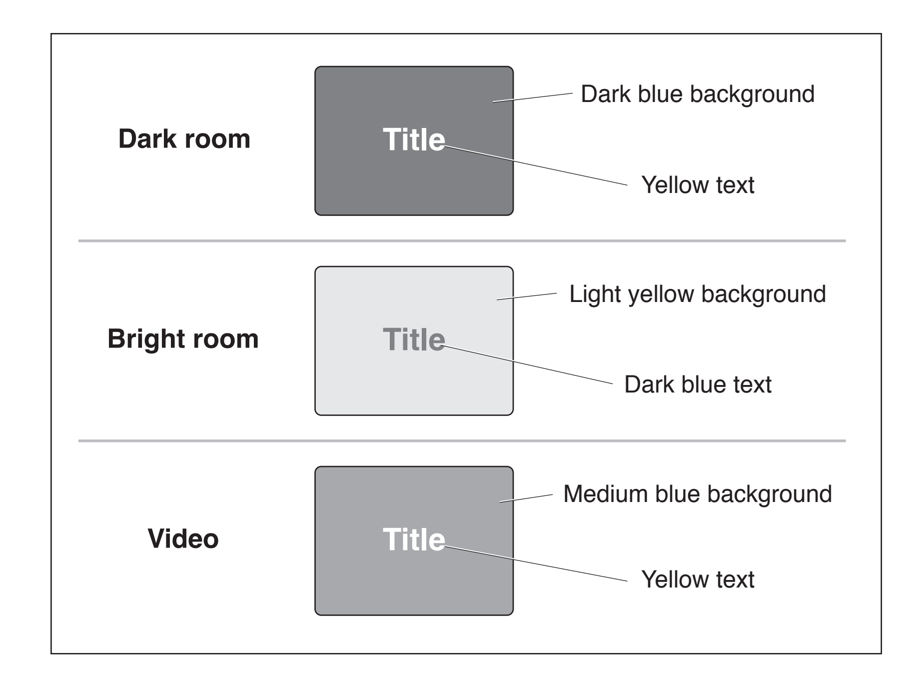


Figure 14.5 *Here are some rules of thumb for selecting colors for different presentation media.*

Subduing and highlighting aren't usually done if the range of usable color tones is so narrow that all of them must be light-shaded.



Using overhead transparencies is something of a rarity these days, but in very large rooms it might be the most effective means of display. The presenter can cover portions of text with a blank card (called a *slip card*), sliding the card downward to reveal each line or section of text as the discussion progresses.

Returning to the topic of contrast, graphics produced for broadcast video require relatively narrow contrast ranges as compared with computer-graphic displays, or film or paper media.

The mood or content of the presentation is another consideration in choosing a background color. A deep blue is by far the most popular color for the backgrounds of business slides. There are two reasons for this:

- Blue objects tend to recede, or stay in the background, in relation to other colors. Blue contrasts well and does not detract from other elements of the composition.
- Blue is the most popular color for corporate logos. This is a content issue: A company audience is likely to be biased in favor of their "school colors" as an overall theme for all slides in the presentation. Yellow and white for titles and text provide good contrast with blue.

Other effective background colors are green and dark gray. Both of these colors are easy on the eyes and compose well with the colors of other elements.



At least for Western audiences, avoid color combinations that are traditional for holidays—red and green for Christmas, blue and white for Chanukah, orange and black for Halloween, yellow, pink, and pastel blue for Easter. This rule is analogous to avoiding familiar tunes in the sound tracks of recorded presentations: well-known color or music schemes will evoke feelings in your audience that have nothing to do with your message.

Pastel colors are often used for the backgrounds of presentations in fully lit rooms, as well as on Web pages. Yellow, pink (or light magenta), light green, and light blue are good choices. Titles and text might be dark blue or black.

Compatibility with other output media should be your final consideration in choosing a color palette. Converting a presentation designed for a display screen to paper handouts, for example, usually involves some redesign for reproduction in a different medium.

In preparing to print on paper, you might be limited to the *grayscale*—or range of monochrome shades—of black-and-white printing. For example, the tones of red and green that you used might both translate to the same shade of gray. To differentiate the colors in the monochrome version, you might have to use patterns to supplement the shading.



Analog forms of copying your visual materials, such as film duplication, require special attention because each duplication step, or generation, tends to add contrast. For example, slide film originals designed for duplication must have narrower contrast ranges than slides intended for projection. However, these days, most photocopying and many printing services use digital technologies that produce virtually identical copies in terms of color, resolution, and contrast.

How Many Bytes Will It Take?

Color capability of computer graphics systems results from two technical considerations:

- How many levels of intensity can be produced by the display or printer?
- How many data bits are associated with each point, or picture element (pixel), in the image?

Computer displays that use 24 bits can render about 16.7 million distinct colors.

So, from the standpoint of technical capability, the number of choices at hand is not the issue.

Color palettes implemented within computer graphics applications deliberately restrict your color choices to a much smaller number that work well together. You might think that 16 is a reasonable number of selections, but when you consider that you will want several shades—at least bright, dark, and subdued—for each color, 16 might not be quite enough. A workable number seems to be 64.

For Web developers, this notion of preselected color palettes plays a part in Cascading Style Sheets (CSS), templates of styles and color selections that can be applied to pages with a click to change their appearance without the effort of redesigning them.

Colors and Computers

You would think that people could agree on how many colors are enough. As I point out earlier in this chapter, process control engineers who design displays for monitoring nuclear reactors might tell you that

four colors may be too few—and eight too many! Their assumption is that, as a scheme of coding, too many colors are apt to be confusing.

Indeed, some accountants feel this way about business slides: Too many colors are distracting and even somehow dishonest—an attempt to obscure the underlying meaning with dazzle. However, it's been my experience that businesspeople—perhaps because of the influence of advertising and the perception that "color sells"—expect color to be used liberally. Even if you take the conservative position that boardroom graphics should be restricted to, say, eight colors, that begs the question: *Which* eight?

If you want to be able to discriminate among four shades of corporate blue, for example, you must also have that degree of discrimination among the other basic colors. Already, you're way beyond eight color tones.

Choosing Computer Colors

Presumably as an aid to the designer, most computer graphics programs offer predefined styles of presentations, which usually include color palettes (although you might only be aware of the palette as a set of preselected colors). If you let some computer program control the design of your presentations, you might have no reason to learn any more about color selection. But just as soon as a supervisor or client requests a color change in a chart you've made or wants a different color scheme, you need to be concerned about details such as color palette definition, order of color assignment, changing of color attributes, correspondence between colors and monochrome textures, and so on.

There are three basic ways of selecting colors in most charting programs:

- Select a predefined color scheme. The color scheme is controlled by the order of colors in a color palette.
- Change the colors of objects explicitly by assigning them different colors from the palette.

- Mix *custom* colors, which can be used to replace colors in a palette or can be given to specific objects.

How Color Palettes Work

A color palette is an ordered arrangement of a set number of colors (or shades of gray if it is a monochrome palette), as shown in Figure 14.6. (This specific implementation of color palettes was used in the application Freelance Graphics for Windows, a product that has been largely eclipsed by the popularity of Microsoft PowerPoint—which offers Design Templates that contain a limited palette of preselected colors.)

In the example palette, positions of colors in a palette are numbered 1–64 (although the end user need not be concerned with the numbers). This numbered order is the sequence in which the program assigns colors to the different types of graphic objects that you create. Specific assignments of palette color numbers to types of chart objects are shown in Table 14.2.

In effect, chart objects are normally colorless. Instead of having a specific color, each object has a color number. The object will take

Figure 14.6 *In a color palette, sets of preselected colors can be implemented as tables of color-assignment numbers.*

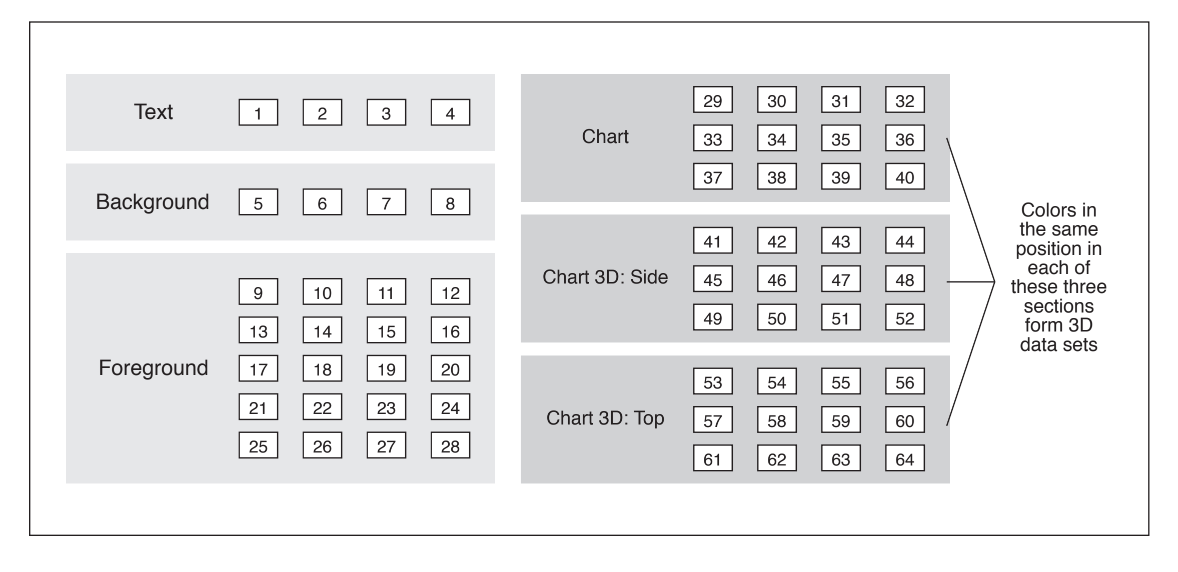
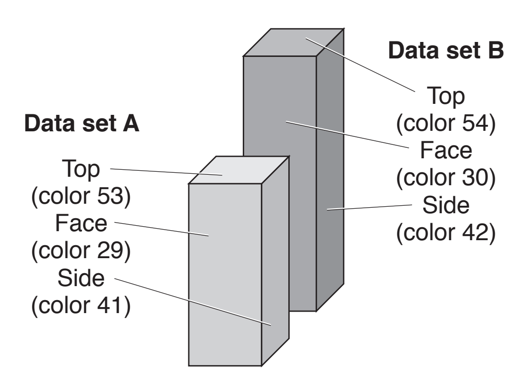


Table 14.2
Assignment of Palette Color Numbers to Object Types

Palette Number	Object Type
T E X T	
1	Titles
2	Subtitles, bulleted text, added text
3	Data chart text
4	Dimmed gray (subdued text)
B A C K G R O U N D	
5	Solid background main color
6	Patterned background secondary color
7	Alternate background main color
8	Alternate background secondary color
F O R E G R O U N D	
9–14	Drawn objects
15	Area and edge of drawn object
16	Chart frame edge
17	Chart frame area (graph background)
18	Organization chart box area or table background
19	3D organization chart box side
20	3D organization chart box bottom
21	Shadow
22	Bullet, line, area
23	Organization chart line
24	3D data chart floor
25	Chart and graph grid
26	Edge (outline in charts)
27	Table border, cell border
28	3D chart wall
C H A R T S	
29–40	Data sets 1–12 (A–L)
C H A R T 3 D : S I D E	
41–52	Data sets 1–12 (A–L)
C H A R T 3 D : T O P	
53–64	Data sets 1–12 (A–L)

on the color or shade of gray found at that position in the current palette. Here, for example, is how the colors of 3D bars are assigned from a palette:



Simply by selecting another palette for use with a presentation, you can change its entire color scheme. Each object in the presentation will take on the color found at the same numbered position in the new palette.

It might be best to think of color palettes as sets of numbered color assignments. Remember, too, that the colors need not be all different. For example, for design purposes, you might want Color 1, used for titles, to be the same as subtitles (Color 2), chart frame edges (Color 16), bullets (Color 22), and table borders (Color 27).

If you change a color in a palette, you change the color of all objects in the presentation that have been assigned that color number. For example, if you change Color 1, which is usually assigned to chart titles, from yellow to white, all of the chart titles in the presentation will be changed from yellow to white.

For any color number within a palette, you can change its color assignment simply by selecting another one of the predefined colors. Or, you can create your own colors by mixing, using either the RGB or the HSL methods.

Mixing Custom Colors

You might think of palette colors as relative and custom colors as absolute. Objects that take their colors from a palette, such as chart elements, will change color when you change palettes. Objects to which you have assigned custom colors, such as drawings, will stay the same even if you change color palettes.



In technical terms, objects that have custom colors are assigned explicit RGB values rather than color palette numbers.

In practice, you might never need to mix custom colors. But you might need to do so if you need to create custom palettes to match the color reproduction characteristics of a specific printer, for instance.

Some computer graphics programs use the RGB method of color mixing. That is, you mix a custom color by making separate adjustments to its red, green, and blue components.

Mixing Colors by the RGB Method

To mix a custom color by the RGB method, you must make separate adjustments for Red, Green, and Blue. The adjustments typically range from 0 through 255 for each of these primary colors. In charting programs, you usually change the settings by typing numbers for the values or by adjusting graphic controls such as sliders (which work like volume controls on a stereo).



If you have a color in mind, it won't be easy to get it by mixing red, green, and blue. The best use of the RGB method is to adjust numeric settings to match values in published tables for printers and other output devices.

Although mixing RGB colors by eye can be cumbersome, here's how to do it if you must: To change hues, add one or two of the component colors. For example, adding some red and some green will increase the yellow in the color you're mixing. Adding blue and green will produce blue-green. Adding blue and red will produce violet.

Increasing red, green, and blue equally will have the effect of lightening, or adding white to, the color. Decreasing them equally will darken the color.

Mixing Colors by the HSL Method

It is easier to mix colors by eye using the HSL method, where you adjust the hue (position in the color spectrum), saturation (tint, or intensity), and lightness (value, or shade). Like the RGB method, you make adjustments by typing numeric values or by adjusting graphic controls.

Changing the Appearance of a Presentation

In most charting programs, you don't control the selection of the color palette directly. Instead, you pick a "look" for the presentation, which is based on a predefined style—a collection of options that includes a color palette.

In computer terms, a style is a collection of *templates*—usually, one for each chart type. You can think of a chart template as a generic description of a particular type of chart, such as a bar graph, which requires only that you supply the data.

Chart templates, in turn, are collections of options. These options include page layouts, as well as object attributes such as edge (outline) colors and fill (solid area) colors, and line styles (solid, dashed, or dotted). When you create a new presentation, the options in the template you choose are applied automatically to all the charts you make.

At some later time, after you have created a presentation, other styles and palettes can be applied to reset many options at, literally, a stroke—much as the appearance of word-processing documents and Web pages can be changed by applying different templates, or *style sheets.*

Viewing Colors on a Computer Screen

The actual colors you see on output can vary considerably from one computer device to another. For example, if you had an old device that could only display 16 colors, only 16 of the palette colors (including black and white) would be shown as solid, or "pure," colors (see Table 14.3). The others would be reproduced by an approximation called *dithering*. (Dithering isn't done much anymore because even

Table 14.3
Pure Colors on 16-Color Computer Displays

Color Name	RGB Values		
	RED	*GREEN*	*BLUE*
Red	255	000	000
Yellow	255	255	000
Olive	129	129	000
Neon Green	000	255	000
Dark Green	000	128	000
Turquoise	000	255	255
Aztec Blue	000	130	128
Flag Blue	000	000	255
Midnight	000	000	128
Hot Pink	255	000	255
Plum Red	128	000	128
Scarlet	129	000	000
White	255	255	255
25% Gray	192	192	192
50% Gray	128	128	128
Black	000	000	000

handheld devices are capable of displaying millions of colors. But this discussion can help you understand the challenges of matching colors between dissimilar display or printing technologies.) Dithering creates a dot pattern that uses two colors to produce the impression of a third. The number of solid colors in the gamuts of other computers are 256, 32,000, or 16 million—quite a range! The more colors in the gamut, the more will be displayed as solid rather than dithered.

Even if colors appear dithered on the screen, they can be reproduced as solid tones on higher-quality output devices.

Those Sexy Chart Backgrounds

A chart background can include all kinds of graphic elements, including not only color, but also drawings or pictures on which charts and text can be overlaid. A background can be a solid color, a patterned blend of two colors, or even a picture. Optional elements would be decorative objects such as borders, symbols, and logos, which often are incorporated in predefined presentation styles (templates).



Liars are fond of using pretty pictures—showing attractive people, showy vehicles, or exotic vacation resorts—as backgrounds for less stimulating, but perhaps much more significant, material. Consider, for example, the timeshare vacation presentation that shows a sunset picture of Puerto Vallarta behind a chart that is supposed to disclose how your condominium investment is expected to appreciate. The audience might be forgiven for remembering more about the sailboats than the sales data!

For best results, the first and second colors of blended backgrounds should be contrasting values of the same color hue, such as light blue and dark blue. When you mix two different colors in a background, some of the gradations of blended color either won't be pleasing to the eye or will be so garish as to take attention away from the objects in the foreground.

How Do You Like This Picture?

Computers typically handle background patterns and pictures as *bitmaps*—arrays of tiny, colored dots. Even though patterns and bitmaps can make attractive backgrounds in color slide presentations, be aware that adding bitmaps can add significantly to the size of the computer disk file in which the presentation is stored.

Changing Background Color

In a charting program, one way to change the background color for all charts in a presentation is to edit the color palette (have another look at Figure 14.6). To change a solid background in the example palette, you would change Color 5 (the first color in the second row). To change a patterned background, you would change both Colors 5 and 6.

Designing for Monochrome Output

A straightforward way to design color presentations for black-and-white or monochrome output—such as paper handouts—is to switch to a palette that has only shades of gray. The program might do this automatically when you select a monochrome output device for printing, but you won't have much control over the result.

Designing charts in color for grayscale output is rather like painting pictures for a person who is color-blind. Remember that you can have very different colors that produce the same grayscale value on paper. Red and green areas of equal intensity (RGB number) will have the same grayscale value and, therefore, will look identical in monochrome. Furthermore, the blue component in any color will be *almost entirely ignored* in the monochrome version.



Liars really don't care whether their color presentations are reproduced well in the black-and-white handouts they leave behind. (That's assuming they have the courtesy to give you any handouts at all!) A really nasty trick would be to show a critical plot in light blue—which will disappear completely when printed in grayscale!

The best solution for presentations that will be reproduced in other media is to use separate color palettes—a full-color palette for slides and a grayscale palette for monochrome printing on paper. You should adjust the shades of the grayscale palette so that there is as much contrast as possible between plots of data sets, for example. If manipulating the grays doesn't result in charts that are absolutely clear, use different patterns, as well.

Designing for Printed Color Output

You might think that a solution to reproducing your color slides faithfully on paper would be to use a color printer. In fact, you must do as much adjusting for color paper output as for monochrome.

Video screens and color film recorders use *additive* color mixing. The result of adding red, green, and blue—all at full intensity—is white.

Printing on paper and exposure of photographic films rely on a process of *subtractive* color, based on how light is reflected through colored dyes, such as printing inks. The RGB and HSL color mixing methods can't be used directly with subtractive colors. When you design presentations for color printing, you must use the *CMY* method, named for its component colors—cyan, magenta, and yellow. Color printing involves layering patterns of ink in these colors on white paper, which reflects light back through them to create color mixtures. However, the combination of cyan, magenta, and yellow at full intensity does not produce a true black (as RGB does with each component at zero intensity). To make up for this, full-color printing uses a fourth layer—black, designated *K* to avoid confusion with blue. The whole scheme is called *four-color process,* or *CMYK.*

You don't have to worry about the technicalities of translating among RGB, HSL, and CMYK. Many charting programs will do this for you. Another guide is the Pantone Matching System, or just *PMS Colors.* A long-standing vendor of specialty art papers and printing inks, Pantone, Inc., has developed a system that establishes equivalents

between standardized color names (such as "Sangria" for a particular tone of red-violet) and their different numeric descriptions in RGB, HSL, and CMYK.

Beyond the matching of colors, be sensitive to differences in color-palette requirements for putting charts on paper. In general, the rules for designing overhead transparencies will be more appropriate than those for slides: Use light-shaded, neutral, or pastel backgrounds and dark text.

For Extra Credit: Using the Color Wheel

Yes, choosing attractive color combinations for your charts is largely a matter of taste. But by now you should be aware of many other important factors that, at the very least, might overrule your selections.

A device for making attractive color choices is the *color wheel*, shown in Figure 14.7, which you may recognize from art class in elementary school.

The color wheel shows how the primary hues of red, green, and blue can be blended to produce the other hues. The blends of the primary

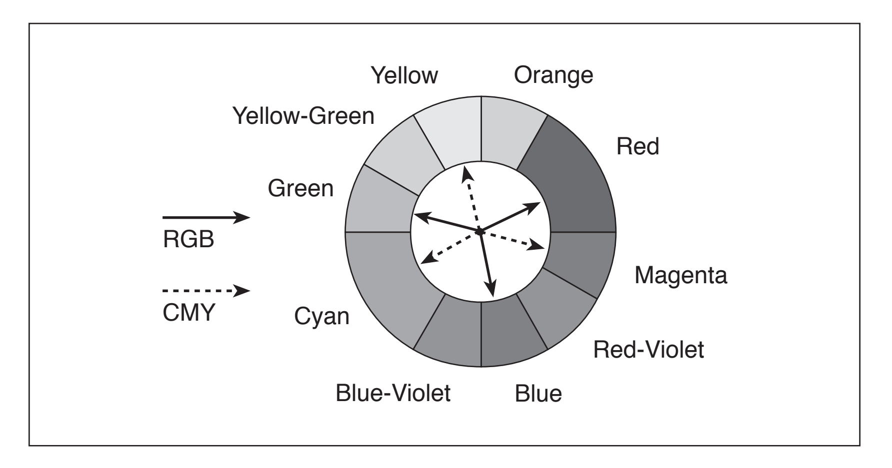
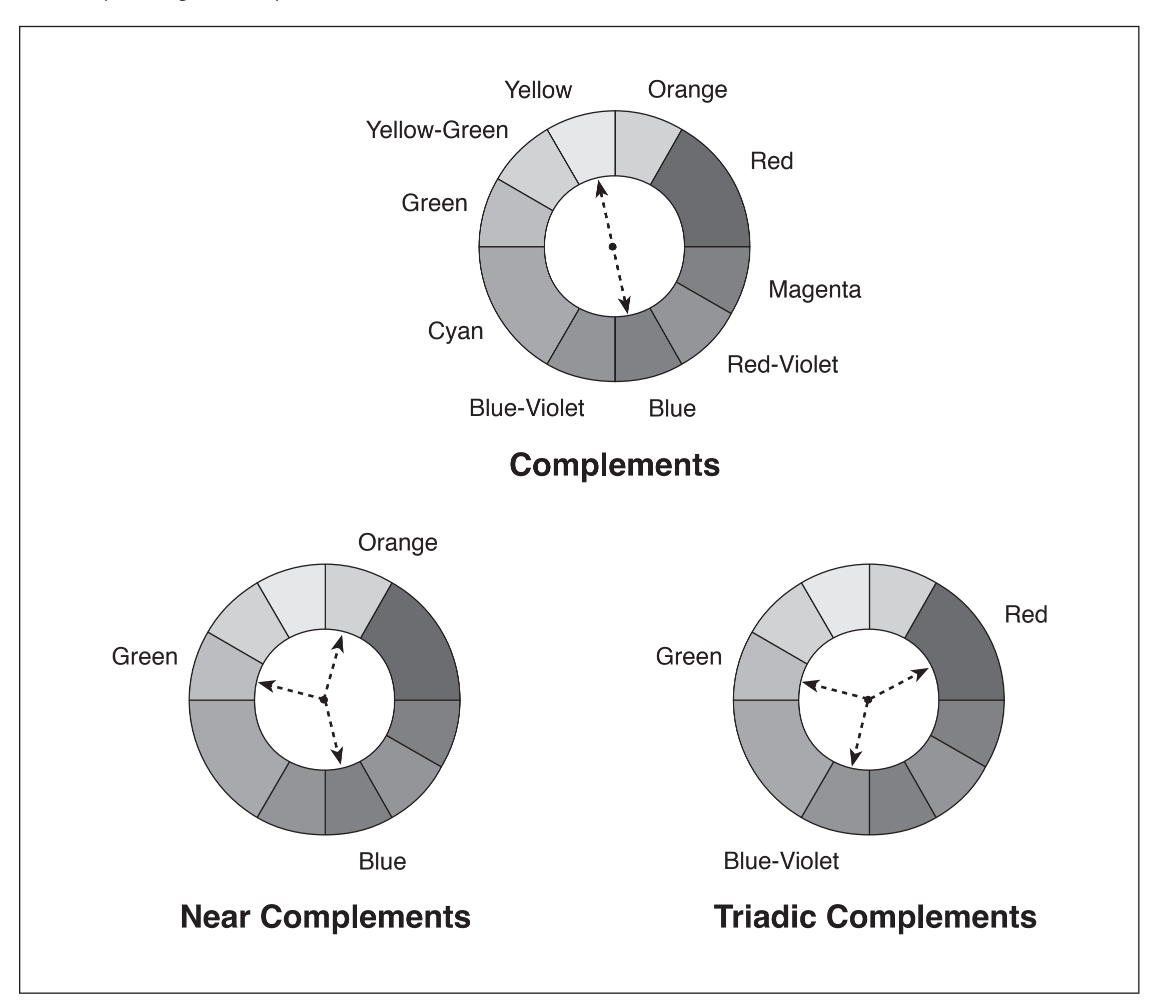


Figure 14.7 *Remember the color wheel from art class?*

hues are called *secondary* hues, such as yellow-orange. (Note that the primary colors of the RGB color wheel are the secondary colors of the CMY wheel, and vice versa.)

According to color theory, pleasing combinations of hues result from certain positional relationships on the wheel, as shown in Figure 14.8.

Figure 14.8 *Positional relationships of colors on the wheel produce combinations that are pleasing to the eye.*



By color-wheel position, there are three types of color combinations: complementary (opposite on the wheel, shown on top), near complements (left), and triadic complements (right).

Each of the predefined color palettes in charting programs is set up according to one of these three methods, based on the main selection of the chart background color. If you don't have a readymade palette to choose from, you now know what the professional designers know about putting colors together!

Getting Smart About Color—A Case Study

In one of my previous career incarnations, I was creative director for a company that produced business presentations. I offer the following true story as a cautionary tale about the uses and abuses of color.

My assignment was to produce a slide show for an audiovisual agency in Detroit. The agency's client was an automobile manufacturer—an influential customer, especially in that town. The slide presentation included word slides, charts, and, of course, stunning photographs of the new car model.

For the computer-generated slide graphics, the client chose backgrounds of deep burgundy. When the original slides were projected, they were gorgeous, in my humble opinion. The burgundy conveyed richness and luxury. I was sure that the client would be pleased.

The agency, which had contracted elsewhere for the car photos, made the decision to handle its own slide duplication through one of the professional film laboratories. (Yes, this was back in the days when presentations were recorded on color slide film.)

When the slide copies came back from the lab, I received an angry call from the agency. The beautiful burgundy backgrounds had turned an unattractive brown—and the film laboratory insisted there was nothing to be done!

I had another look at both the originals and the copies, then paid a visit to the film laboratory. I knew that the overall tint of the film copies could be adjusted to more closely match the computer-generated colors. I had a long and heated conversation with the technical chief, who explained, "People in this town look at three things in a film print—people, sky, and cars. When we adjust the tint, we go first for the faces. And auto paint is very difficult to reproduce, especially the metallics. We have to get the paint job to look just right without turning the people green. The graphics just aren't my worry. Who else but you knows what color they were to start with?"

The situation had been made worse by the selection of the exceptionally dark burgundy background color, which was nice for projection but could not be rendered properly by the duplicating film without sacrificing some other range of colors in the rest of the presentation. Unfortunately, that hearty burgundy now looked more like burnt pot roast. But, if we were to adjust the tint of the copies to enhance the background color, we'd ruin the live photography, including the color of the client's product—that shiny, new car.

Through my previous long and painful experience dealing with the auto makers, I had learned the consequences of getting the car color even slightly wrong. When one of the executives saw some film prints I had adjusted for color, he demanded, "Where did you get that vehicle? We don't even make that color!"

The solution to what I'll call "The Case of the Muddy Burgundy" was to make a special set of slide originals—with lighter backgrounds in the same hue—for duplication. (Again, digital reproduction methods, for the most part, don't suffer these problems.)

So, consider that a single photograph of a client's product in a slide show could very well dictate the colors of the entire presentation—that is, if you aim to show the truth, as the client sees it.

And that's not just a matter of taste!

15

Social Media

Or, The Disinformation Superhighway

When I wrote the first edition of this book in 1994, the idea of deliberate manipulation of graphics to deceive seemed remote. In my years of supervising the design and production of management-level business presentations, I never saw any evidence of deliberate fraud. But what I did see, almost on a daily basis, were unintentional errors, omissions, and mistaken design decisions. Those mistakes probably resulted in misinformation for and misperceptions by our audiences. And those mistaken impressions, in turn, may have resulted in poor business decisions. (But at my management level, I was not aware of those consequences.)

Now in the second decade of the 21st century, the game has changed fundamentally. Our vocabulary has a new term for outright lies—*fake news.*

Fake news exists because it can work to change opinions—or reinforce existing viewpoints. But, in many ways, the effectiveness of fake news defies logic. As light overcomes darkness in a flash, substantiated truth should extinguish lies. However, the effectiveness of deliberate misinformation—or *disinformation*—has nothing to do with logic.



Fake news works because it leverages emotion.

Hollywood film producers will tell you that, when an audience leaves the theater, they will soon forget all but the most sensational lines of

dialogue. But they will remember vividly the emotions with which the actors delivered those lines. And the impressions can persist in a viewer's subconscious mind for a lifetime. That's why, in the movie industry, writers are *important*—but movie stars are supremely *influential*. A star is a star because he or she exudes credibility. The force of their on-screen personalities is an irresistible emotional pull. Audiences believe what those stars say on the screen, not because their speeches necessarily make sense, but because they invest their words with emotional power.

In screenwriter William Goldman's book *Adventures in the Screen Trade,* he advises that movie stars don't like to deliver long speeches. They dislike playing characters who explain. Ideally, the star wants to listen to a supporting actor describe the plan. Then the star gets to inspire and motivate the plan with emotion by ordering, "Let's do it!"



If a news item or an opinion expressed in social media (or on any platform) leverages an emotion, the emotion—and not the underlying facts—will fuel its distribution and acceptance.

One of the inspirations for this book was the 1954 classic by Darrell Huff and Irving Geis, *How to Lie with Statistics.* Chapter 1 of my book, therefore, focuses on statistics and data manipulation. The premise is that charts are no better than the data from which they are generated. And some forms of data manipulation—even when the results produce distortions—are not only easy to do but can also seem eminently logical and convincing.

The principles described in this book can be applied to all kinds of illustrations for online posts. But the problem isn't exclusively graphical. In the same way that skewed statistics might underlie bad graphics, skewed logic can lurk beneath fake news. So, by applying the same skeptical sensibilities I brought to the earlier chapters, I describe here some techniques of misinformation—and, yes, deception—I've observed in social media.

Now, can I safely assume you are an avid and routine user of social media yourself? You therefore might think my observations here are obvious to everyone who has any online experience at all. These tips might seem self-evident and commonsensical.

However, if everyone knows how the game can be manipulated, why do so many people keep forwarding bogus or flawed messages to all their fans and followers?

First off, an appeal to forward a message to *all your friends* should be a red flag in itself. Moreover, pitches for seemingly worthwhile charities or pictures of kitties or dire warnings about hackers should be just as suspect as biased political messages. Why? Because a basic operating rule—which has been true since the earliest days of the Internet—is:

THE AUTHOR OF A MESSAGE HAS THE RIGHT TO TRACK IT EVERYWHERE IT GOES.

On average, the everyday user might not know how to do that. But the administrator of an email server who starts a message thread assuredly does. Ergo, if you do share a message with lots of your friends, the instigator of that thread can vacuum up all those identities. Depending on the platform used, the author of the message might also be able to get your friends' personal contact information, even their profiles. Hackers can profit from such unethical behavior by selling those lists to third parties.

With charts as well as with statistics, most of the ways you can distort information aren't particularly sophisticated. The complexity, mathematics, or geometry involved don't require advanced training or skills. Quite the opposite—mistakes can occur because we're prone to ignore underlying assumptions that are so basic we don't think about them. For example, harking back to our discussion of pie charts, it might seem obvious that making a circle twice as tall will indicate that its total value has doubled. As shown in Figure 15.1, the *area* of a square drawn around the larger pie is mathematically *four* times the original height. It's the *square* of the pie's doubled diameter (the math is easier with squares than with circles).

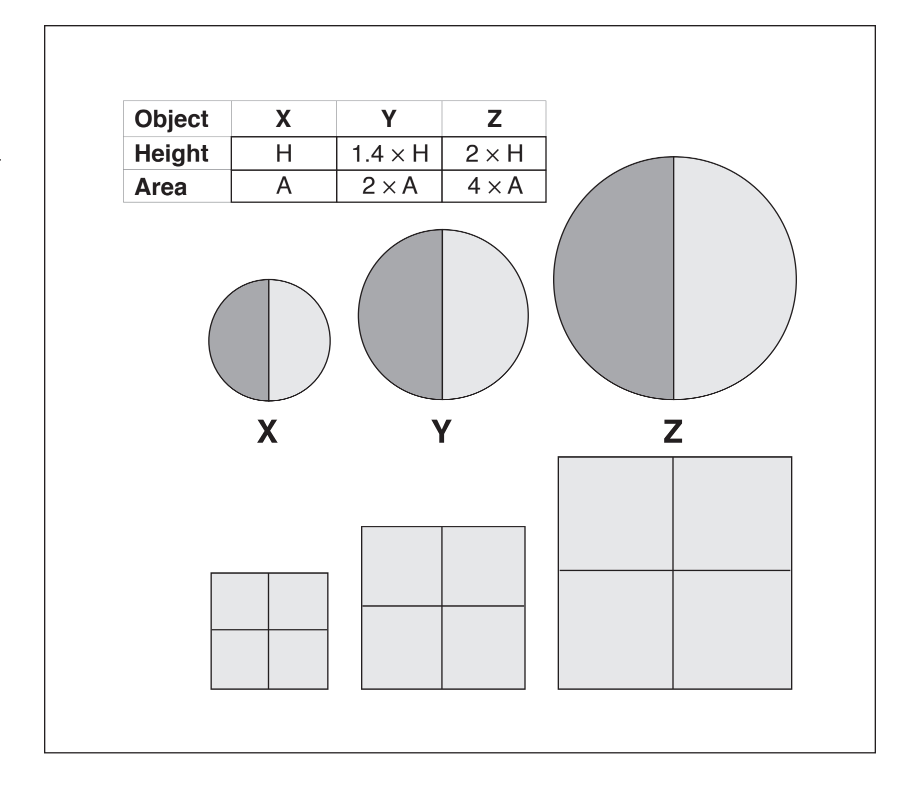


Figure 15.1 *As explained in the discussion of pie charts in Chapter 2, doubling the height of a pie chart actually increases its apparent area (the size of the square around it) four-fold.*

What's more, comparing the areas of pie charts is fallacious in itself. Pies are percentages, indicating the relative angles of the wedges as shares of the pie. The absolute area of a pie is meaningless and always stands for one and only one quantity—*one hundred percent* of the pie.

Similarly, disinformation in social media can rely on assumptions that seem commonsensical. Here is where leveraging emotion clouds good judgment. On reading a social-media snippet that evokes strong feelings, our analytical attitude can give way to warmth or fear or anger. The impression is all the more powerful because the emotion hits us first, before we apply logical and critical analysis. Like those theater patrons, we remain vulnerable to our first emotional impressions, despite our later efforts to reason them away.

With that thought in mind, I see these techniques in use every day in posts from suspicious as well as from responsible news sources. In some

cases, the mistakes might not be deliberate. Especially in the early days of the Internet, we heard stories of reputable news organizations that skipped fact-checking in a rush to post a story before blogger sites that did little or no background research could break the news. Although commercial news outlets still talk a lot about journalistic responsibility, the pressures on rival players in the blogosphere to break exclusive stories have only increased.

How to Abuse Social Media

Misleading posts may use the following techniques. I encourage you to be watchful and think of more. (And please call them out in your own posts whenever you see evidence of fake news.)

1. Change the Order of Events

This technique can work in either a breaking story—when the events are still under-reported—or in recounting a historical episode—when the story has lost its freshness and readers might not remember the details. In its simplest form, an unscrupulous author could reverse the occurrence of events, thus confusing—or even reversing—cause and effect. Doing so can recast the perpetrator as the victim, and vice versa.

For example, in the original version of a story, a state legislature passed a bill to regulate a hazardous chemical. Then the manufacturer changed the formulation of its product to exclude the offending ingredient. A mistaken news account might first state that the manufacturer changed the formula, then add that lawmakers banned the omitted ingredient. In the true version, the manufacturer's role seems reactive. The company might have known the chemical was hazardous but was only forced to change its product when the use of the chemical became illegal. In the second, misstated version of events, the manufacturer might be perceived as prescient and pioneering. Also, the elected representatives might look like slow reactors who are papering over their longstanding failure to protect the public. A little research would establish which

version is correct. But some readers may have a bias either toward free enterprise or toward regulatory reform. And the emotions behind those biases could well determine which version of events the reader is likely to believe.

Changing the order of events is a basic (and usually legitimate) tool of video editing. Chronology is important. I'm not saying that all biographies should begin with "I was born," but be aware of shifts in occurrence, particularly if the edit changes cause-and-effect relationships.

2. Accuse the Accuser (or Scoop the Scooper)

In a variation of the above event-swapping technique, a guilty party strives to be first to break the news of its own crime. The wrongdoer blames someone else—typically, its opponent—in the same news release. This is an age-old political dirty trick. The crime may be real, and so is public anger about it. However, until the story breaks, the anger may have no focus. There isn't yet a consensus about whom to blame. When the story breaks, public ire finds a target, and it can be difficult for the unfairly accused parties to refocus attention on the facts.

Take as an example a variation on the above story about a banned chemical. The public has known for years that a stream is heavily polluted. During the run-up to a local election, the opposition party accuses the incumbent party of failure to address the problem. The accusation omits the core fact that the opposition party has been consistently blocking the corrective legislation for years. Again, because of bias and emotion, proponents of the opposition—as well as angry fence-sitters—are likely to accept the story's accusations at face value. At the same time, members of the incumbent party and their sympathizers may be motivated to dig into the history of the problem to find out why the regulatory measures were never passed. But their efforts to publicize those facts will be at a disadvantage because the opposition has broken the story first. The liars have successfully invoked the ire of the electorate—igniting the story's emotional fuel—at a critical time.

3. Stretch the Analogy

As any student of elementary logic knows, analogies can be inherently fallacious:

A implies B
and
B implies C
but it does not necessarily follow that
A implies C

The Venn diagram in Figure 15.2 shows why analogies can be false. The Set A shares some elements of Set B. Set B shares some elements with Set C. But despite their separate affinities with Set B, Sets A and C do not share any elements.

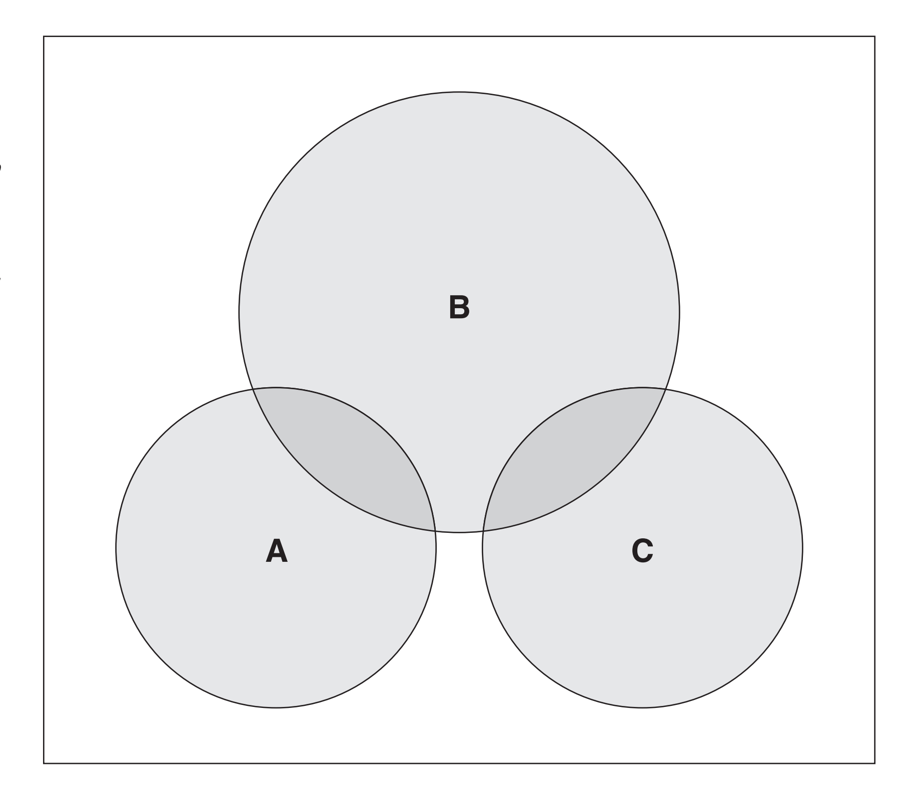


Figure 15.2 *As this Venn diagram shows, some elements of Set A are contained in Set B. Some elements of Set B are contained in Set C. But Sets A and C have no elements in common.*

To see how silly these types of messages can get, consider this example:

My political opponent Jack has red hair.

Jack's ally Lucille also has red hair.

A man with red hair robbed the supermarket last night.

We must monitor the behavior of red-haired people more closely.

This red-haired paranoia is a jumble of confused analogy and guilt by association. And when you substitute skin color, religion, or political affiliation for hair color, the examples go from silly to seriously wrong.

The practice of argument by analogy won't go away because its logic is flawed. Even though analogies shouldn't be cited as proofs, some analogies *are* valid. In the Venn diagram, if the facts were different, some or all of the elements of Set C *might* be contained in A. An analogy on that basis describing the commonalities of A, B, and C could therefore be useful as an illustration or an example. And I think that's why we are all prone to use analogies in casual conversation. Doing so helps us explain the unfamiliar by describing the familiar.

But, from a logical viewpoint, a colorful illustration is not a proof. Even courtroom attorneys who should know better might attempt to sway a jury with a vivid example based on an analogy:

"Counsel for the defense would have you believe that the accused drew a weapon because she felt threatened. But why was she carrying a weapon in the first place? Do you carry a handgun every time you enter a supermarket?"

In this hypothetical case, the prosecutor is making an analogy between the routine personal behavior of the defendant and that of the juror.

Jurors might legitimately question the accused's reasons for carrying a gun on a particular day. But, without more information, assuming that their mental states should be the same as the accused's on entering a store is wrong.

One reason stretching an analogy can work is because it diverts attention from the argument's lack of logic. A portion of the message is undeniably true and familiar, which lends credence to a fallacious conclusion. In other words, it leverages a preexisting belief. And that belief persists because it is invested with emotion, typically anger.

4. Focus on a Side Issue

Focusing on a side issue also diverts attention, not by overruling an analogy's lack of logic, but by stretching its dubious conclusions even further. In such messages, the content may only be peripherally related to the problem at hand. But the emotion surrounding the side issue is strong. By finding a way to link the side issue to the problem, however remotely, the messenger stretches the analogy so far as to include the unrelated issue. Doing so overwhelms and diverts attention from the core issue.

The effect of this deception can be all the more potent if the side issue seems worthy of itself. People are right to be angry about it. But they may be wrong at this time to take their eyes off a more significant problem or a more serious threat.



Use fiery emotion to fuel a different fire.

A blatant everyday example in social media is the misuse of hashtags. If a topic is hot and trending, it can be unethical and misleading to append the tag to a message that has little or nothing to do with the trending topic:

> Elvis has left the building, but I'm still writing books. #elvis #myhashtag

The crucial element here is timing. The trending hashtag won't have the same attractive power when its topic is not trending. Similarly, releasing

two unrelated but breaking stories at nearly the same time can create an association between them. Even more disruptive, releasing one soon after the other might give the mistaken impression that the first is the cause and the second its effect.

5. Release Late Friday (or After Closing Bell)

A simple but often effective trick is to release bad or embarrassing news late on a Friday when people may be leaving work early or departing on weekend trips. The same technique is used routinely on Wall Street to avoid stock-price panics by releasing disappointing news after markets have closed.

The old barroom saying, “It’s always five o’clock somewhere” reminds us that, in today’s global information marketplace, tricks of timing aren’t likely to be effective unless the primary negative impact to be avoided is local. But in a hotbed locale, such as Washington, DC, or the New York financial district, mitigating the reaction of locals might still be worthwhile.

For favorable news—and for product promotions—the timing of releases can be engineered to advantage. For consumers and business audiences alike, Sunday evening has more users online than at any other time. And to reach business audiences during the week, releases can be scheduled either during “ablution time” (before breakfast) or right before the lunch break. Software apps that handle automated postings can be set up to generate a *rolling release* schedule—by which the release time for each region is synced with its local time zone.

6. Confuse with “Alternative Facts”

In the strict, dictionary sense, *facts* have no alternatives. Verifiable statements become facts when repeated and independent verification proves those statements to be irrefutable. While there can be no alternative facts, there most certainly can be a multitude of alternative *interpretations* of the same facts. Here’s the bedrock distinction in journalism: A news story should report facts, and an editorial essay can express interpretations. An editorial can legitimately present opinions, speculation, and even hearsay.

The distinction might seem elementary. And yet there is an intense debate in social media about the extent to which news stories may be written from slanted viewpoints. It may be alleged that opinions and conclusions are being implied between the lines. Sometimes this is done by being selective about the facts presented. Omitting an essential and relevant truth can be as damaging as a lie.

But the danger of having alternative versions of the truth is much broader in scope than just the veracity of individual news items. A time-honored technique of disinformation purveyors is to flood the information marketplace with competing versions of the same story. For example, the press agents of a movie star might learn that a journalist is working on an unauthorized biography of the celebrity. That book will break a major, embarrassing scandal. The savvy publicist knows that, in these days of instant communication, the truth can't be easily suppressed or covered up. So the counter-strategy will be to commission several other books, each with different opinions about what happened. As well, news leaks, interviews, and planted feature stories—each with conflicting claims—can be created and timed to coincide with the release of the offending book.

In such strategies, the more sensational the conflicting claims, the better. Why not create unsubstantiated conspiracy theories about visitations by space aliens? The result can be guilt by association: The inherent lack of credibility in some of the stories can rub off on the more responsible ones. Thus, the believability of any of the stories—including the truthful ones—can be refuted with, "It's just another of those crackpot theories."



If you can't extinguish the truth, smother it in noise.

7. Point to Anomalies as Trends

Is one inconsistent data point in a series a mere blip or might it be a significant *outlier?* In building a message around this fact, it's again a question of interpretation. The ongoing global debate about climate change presents us with such examples daily. One instance of "the hottest day on record" is not, in itself, evidence of a warming trend. But several years of days that exceeded records could well be.

There's much more about metadata analysis and long-term trends in the next chapter, including some startling information about the climate-change debate.

8. Use a Pitchman

Whether in television spots or at business meetings, a time-honored presentation technique poses a pitchman with a pointer in front of a flip-chart or poster. The intent may be for the speaker to highlight significant portions of the graphic. However, there's a fundamental principle of human nature at work here we don't often think about. That is, the facial expressions and the body language of the pitchman will always be more interesting than the graphic.



A pitchman with a persuasive personality could make you believe the opposite of what you would see on the chart if you studied it instead.

9. Animate Your Charts

A promotion strategy of some social-media platforms is to give wider distribution and better search rankings to posts with video as compared to still photos and graphics. But one essential characteristic of a chart is that it is meant to be studied. Animating the chart may add visual interest and draw audience attention. But doing so might also significantly reduce study time and attention span. Yes, you could counter this effect by putting the video in a loop, but viewers might find the repetition annoying.



Making the viewing experience more transient will diminish the understandability of a chart.

10. Use Nonstandard Chart Formats

A virtue of standard chart formats is their familiarity. Viewers know pies are for percentages and *xy* charts typically plot quantity or value versus time. However, website editors may want eye-catching graphics

to increase hits on and sharing of posts. Or designers may want to be innovative to impress their clients and peers.



A nonstandard chart format will inevitably require much more study time from the user. But attention spans are short and getting shorter.

11. Abuse the Net Promoter Score

These days it seems every time you check your email or visit a vendor website or call customer service, you will be asked to take a customer satisfaction survey. Veteran marketers understand what's happening here, but many consumers don't.

Customer surveys are concerned primarily with the value of the vendor's brand, not necessarily the actual quality of its product or service. Aren't those the same thing? Or so strongly related you can't separate them? Not at all. The diagram in Figure 15.3 shows how most people assume the survey process works.

It's natural to assume the survey will be used to improve quality.

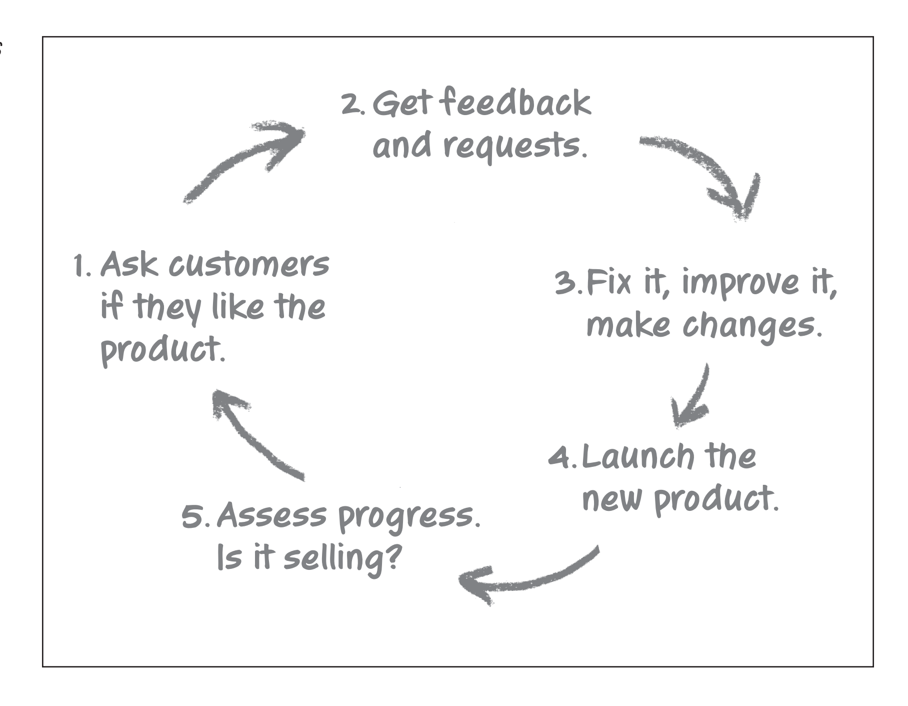


Figure 15.3 *Consumers who receive survey forms may assume their responses will trigger improvements in the product or service.*

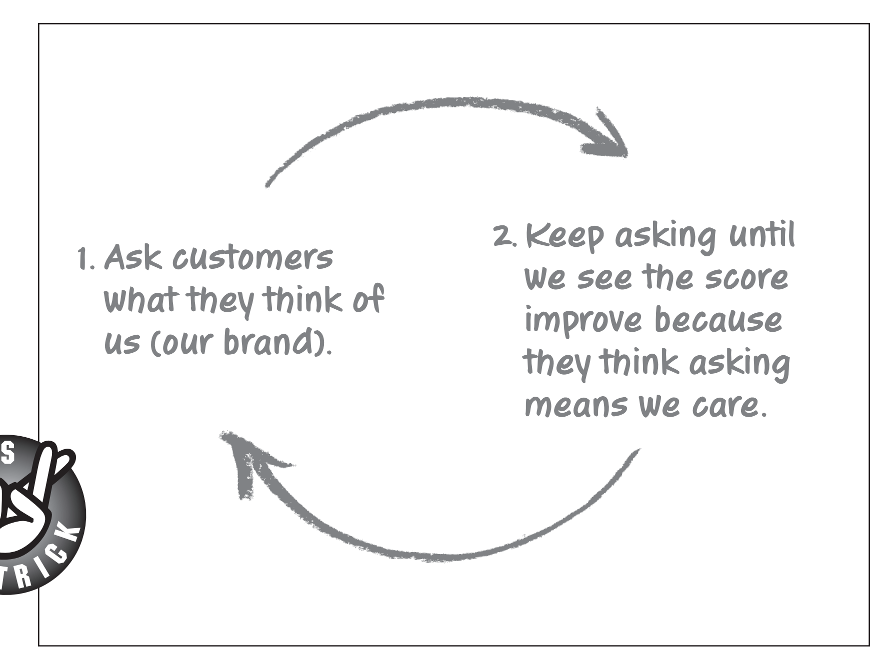


Figure 15.4 *An under-appreciated effect of surveys can be that merely asking the question can cause favorability ratings to increase. If you're a top manager focused on near-term results, why not spend your limited budget on quick, cheap surveys rather than on time-consuming, expensive product development?*

But in practice, there may be only two steps, as shown in Figure 15.4.

In this shortcut process, the *survey itself* is the only customer service being rendered. It's nothing more than a positive public-relations message. It's hype.

Lest you think I'm a curmudgeon on a rant, consider the metrics behind these surveys. The primary methodology in use these days is called the *Net Promoter Score (NPS)*.

Appropriately used, the NPS measures the value of a brand but not necessarily the quality of the product or service. Repeating NPS measurement in iterative surveys measures changes in public perception of the brand over time. One use of NPS might be to put a financial assessment on a brand in preparation for a company merger or acquisition. In such cases, analysts may be legitimately more interested in perceptions than actual quality. For example, NPS could be useful in estimating the market value of a startup company that has received widespread publicity but has yet to launch its first product.

Even though questions posed in a survey may ask about quality, the NPS metric focuses primarily on the customer's willingness to endorse. The core question is: *On a scale from 1 to 10, how likely are you to recommend our brand to someone else?* (That's the *promoter* part.)

However, the functional department in the company that is charged with conducting the surveys—typically, marketing—may have little or no connection or communication with the functional area responsible for improving and delivering goods and services.



Simply asking the question tends to improve the favorability of the brand. The survey itself strongly suggests that the company cares about quality. However, over the short term, it may be more cost-effective to spend money on more surveys rather than on quality-improvement activities.

There's a fundamental disconnect here between functional departments. The marketing function is focused primarily on communication—developing and distributing messages that promote the company and its brands. Sales strategy and competitive pricing may also be part of their responsibilities. But they rarely have the span of authority to order changes in product design or customer service. Marketers must flow suggestions for product development upward to top management. And as long as the value of the brand seems to be improving in the public mind, management might think marketing is doing its job—as well as assume that the product or service needs no improvement.

Improving quality by upgrading the product or training personnel better is not only expensive but also takes time. Management may be focused on the near-term value of the company's brands because that's all most investors care about. The perception of value is value, or so the wise market says. Now, you might say, true value eventually will win out. And to that I say, no one is looking that far ahead. Today's stock market price and next quarter's earnings report are about as far ahead as company management and investors may be motivated to look.

Making Real Quality Improvements to Customer Service

To significantly improve customer service, an organization could:

- Take the budget spent on too-frequent surveys and apply it to quality metrics that measure how well the product or service performs (e.g., Mean Time Between Failures [MTBF], time to resolve) and process improvement.
- Use fewer phone robots and third-party call centers. Staff customer-service functions with live people who have *customer intimacy*—that is, who live in the same region as the customer, use the same products, are on about the same economic level, and so on.
- Train those people thoroughly before allowing them to deal with customers.
- Retrain, retrain, retrain to keep them current on the changing marketplace and the company's continuously upgraded offerings.
- Incentivize, rather than browbeat, employees to care about their customers.
- Retain and reward employees when quality-assessment (and, yes, the survey) numbers improve.

The Impact of Metadata and Big Data

Metadata is data that summarizes other data. As the term is commonly used these days, it refers to the large-scale, long-term process of data reduction, or global summary. *Big Data* is the technology and the industry that delivers all kinds of large-scale metadata.

I'll explore both concepts, including their processes and potential for abuse, in the next chapter.

16 Metadata and Big Data

Or, A Pointillist's View of the Big Picture

Metadata and Big Data have been hot topics in recent years. There is an ongoing worldwide debate about the ethics of widespread data collection, the potential for invasion of individual privacy, and responsibilities for data custody and disclosure.

Metadata, of itself, is not at all about individuals. Metadata encompasses all forms of data reduction. Put another way, metadata is a generalization, an aggregate of specifics. In a metadata analysis, thousands, perhaps millions, of individual case studies may be collapsed into a few data points. Only the large-scope view—only the big picture—forms the basis for conclusions.

Metadata is a high-level summary of data from more detailed records. The term is commonly used to describe the large-scale, long-term collection of data based on discrete data samples from demographic populations. The data samples by themselves are unimportant to the overall result. But the generalizations from millions of samples might yield new and valuable information.

Big Data is the technology and the industry occupied with collecting, aggregating, analyzing, and storing all kinds of large-scale metadata.

Public Concerns About Metadata

Basic concerns about metadata collection and its use include:

- How is the custody of individual and transactional data safeguarded? Even though metadata generalizes about populations and not about individuals, is the underlying data stored in a way that could be traced back to those individuals? If so, how might that data be used or exploited against the best interest of the person?

  For example, knowing how many people within a large demographic sample are diagnosed with diabetes every year could contribute toward a better understanding of whether a particular preventive protocol is having any effect. However, the association of your name and clinical record with the disease should be private, something you alone decide to share with caregivers, family, and employers.

  At the level of metadata analysis, the impact of any individual record is negligible. There should be no reason for the analysts to know who you are.

- If individual transaction data is to be retained, there is another worrisome risk: Applying the generalizations produced by metadata analysis to individuals or to specific cases is likely to be inaccurate or downright wrong. And, as examples in this chapter show, making any generalizations from transaction-level data is counter to the objectives of metadata analysis and is likely to be misleading or incorrect.

  For example, a generalized conclusion that people over the age of 65 are more likely to be at fault in highway accidents does not, in itself, justify revoking your license on your 65th birthday.

The Quest for Metadata Correlations

I am not an expert on metadata collection or analysis. However, in my research for this book, I discovered an extremely helpful and insightful text, the 2015 whitepaper by Hans-Martin Stoenner, *Correlation of Data Series: A Scientific Study on the Selection of Meaningful Variables and Functions for the Separation of Trends, Cyclic Parts and Scatter from Data Series.* This study is brief but dense with advanced math. I don't pretend to understand all of the book's implications for emerging disciplines and analytical approaches. But I do recommend it, not only to instructors and to database application developers, but also to serious students. The paper's introduction is nontechnical (I believe I understood it). I summarize some of Stoenner's approaches and conclusions here because they seem to me to be at the core of what metadata analysis is—and *isn't.*



At a basic level, the goal of metadata analysis is to find correlations between data series. From a correlation, the analyst may be able to infer from the data series that it causes the effects in another series.

A correlation exists when the pattern of a given data series closely matches the pattern of a variable being tested. For example, Stoenner wondered why usage of airports all over the globe is increasing from year to year. What variables might be sampled and compared to explain this trend? Plausible candidates might be the growth in worldwide population, the growth in global economic product, the cost of aircraft fuel, the price of tickets, or some other factors. All of these factors may be contributing, but the *root cause*—the cause of all the causes—turns out to be simply population growth.

Although this conclusion might seem obvious, related inferences are not. As Stoenner discovered, the shape of the airport usage curve matches the shape of the population growth curve closely. Thus, population is the only variable that's essential to explaining the trend. All other variables, such as the cost of aircraft fuel, are short-term and have little impact on the overall, net-upward trend.

Component Trends

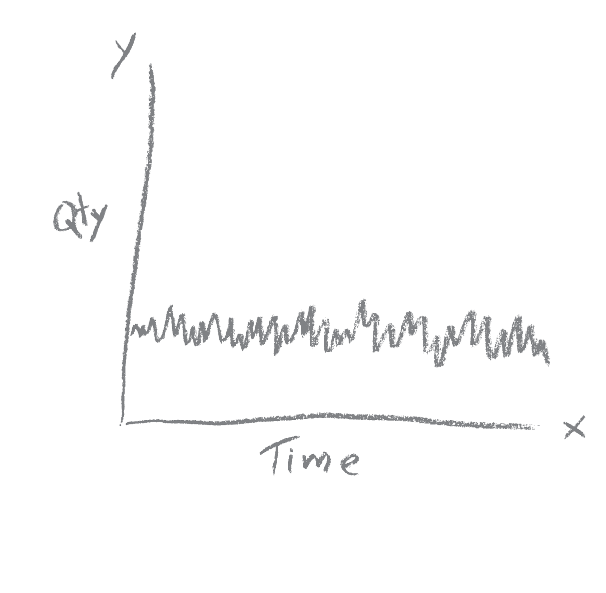
In metadata analysis, variable factors fall into three basic categories:

- Long-term causes and effects
- Drivers of cycles
- Artifacts of volatility

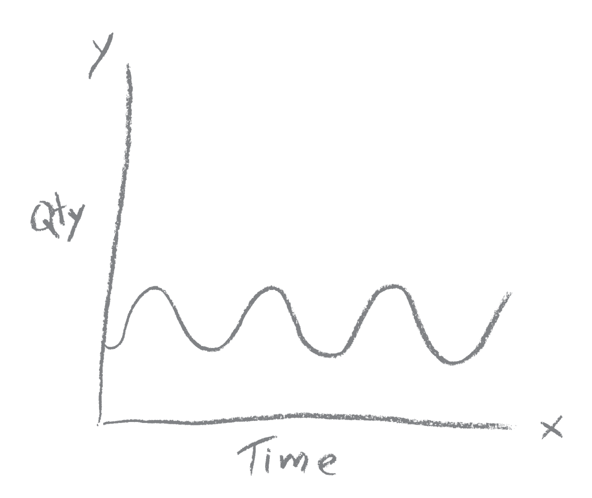
In graphical metadata analysis, a basic goal is to derive a long-term trend from aggregated data for the variable under study. To derive the trend, it's necessary to remove volatility and shorter-term cycles from the picture.

In such charts, the data series are typically time-dependent. That is, the flow of time is inherent in the order of the data points of the series. Any such a time-ordered data series is called a *time series*.

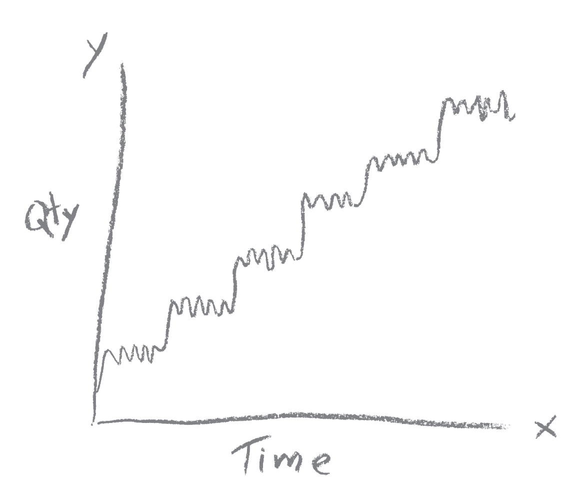
Volatility looks like this:



A cycle looks like this:



An initial plot of metadata—including volatility, cycles, and trend—might look like this:



In this third plot, volatility gives the curve its jagged look. Cycles produce a stair-step effect because as each cycle recurs, the trend line has moved progressively upward.

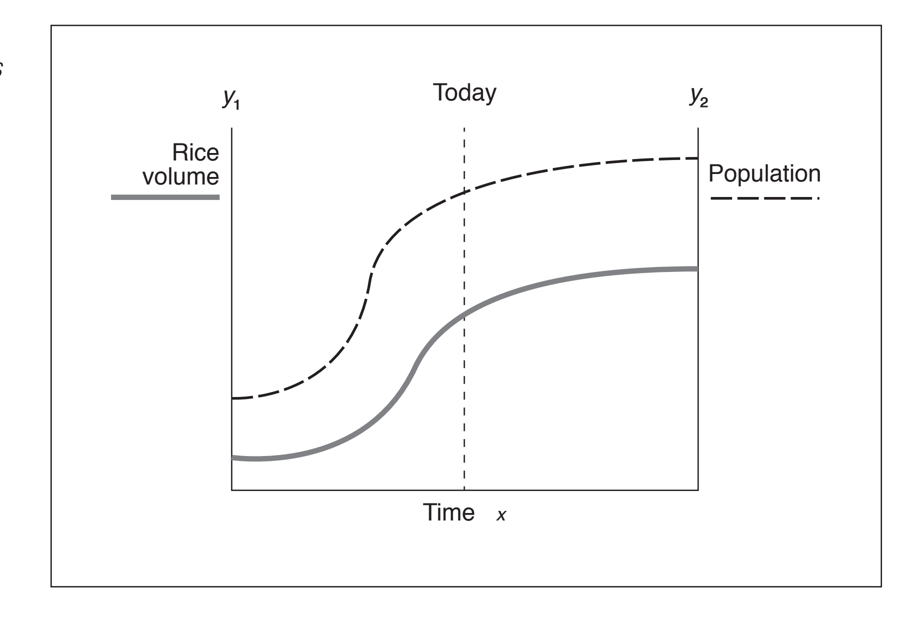


Figure 16.1 *The two plots on this dual-y axis chart compare long-term trends of rice consumption and population.*

Now, here's where things can get interesting. Suppose that the metadata series just presented is a plot of worldwide rice consumption. Removing the short-term data (volatility and cycles) produces in the lower curve in Figure 16.1, which is plotted against the left vertical (y_1) axis. On this dual-*y* axis chart, global population growth is plotted against the right vertical (y_2) axis. In this (admittedly, hypothetical) example, the correlation is obvious—the curve patterns match closely. The volume of rice consumption is proportional to the number of people who must be fed. It's also reasonable to assume that population is the *cause* and rice consumption is the *effect*. The more people we must feed, the more rice gets consumed.

Limits to Growth

Notice that both of the curves start to level off in future years, after the milestone in the chart labeled *Today*. The rice curve is an example I devised after reading Stoenner. (So if it's in error, I alone am to blame.) However, the shape of future population growth is his extrapolation. I won't go into his arguments here, but he shows convincingly that the

number of humans on the planet must peak at about 10.9 billion around the year 2100. However, his math says nothing about *how* population will be limited. That's because all metadata analysis can do based on correlations is to extrapolate *the probability of occurrence.* His forecast may hold, but only if existing conditions do not change. One might subsequently apply metadata analysis to discover which resource limitations (*e.g.,* water, air, land) are most likely to cause population growth to slow down.

Continuing with the rice-consumption example, consider why shorter-term fluctuations are *not* likely to be limiting:

Volatile Factors

The noisiest, or most scattered, data points are short-term variations. For rice consumption, contributing factors at this level might be changes in land prices and fuel costs. But none of these factors is sufficiently powerful or sustained to drive the overall trend.

Cyclic Factors

The stair-step effects in the metadata plot are longer-term changes, which tend to recur. An example would be seasonal weather conditions. Cycles may be significant enough to affect individual consumers, farmers, and marketplaces from one year to the next. But, like shorter-term volatility, the cycles do not arrest the overall trend.

Metadata Analysis of Climate Change

Stoenner goes on to apply his approach to the analysis of climate-change trends. I won't attempt to describe his analytical process here. The result is a correlation pattern similar to those just described for my rice-consumption example. He finds a steep, upward long-term trend in temperature, which is mitigated by the current downward trend of a shorter-term natural cycle. At the present time, the two conflicting causes effectively cancel one another. Thus, it appears the Earth is now in a hiatus period when net temperature increase is fairly flat.

I cite this example because it highlights the dangers of jumping to conclusions from metadata analysis, particularly if short-term considerations lose sight of the long-term trends. Stoenner wisely points out that the hiatus may be used by some observers as evidence for markedly different conclusions—*both* of which may be wrong! That is, because temperature is not currently rising in the short term, one might conclude either that:

- Human policies and control measures must be working.
- The climate-change phenomenon does not exist.

Again, metadata analysis is all about large-scale data collection and long-term projections, both into the past for patterns and extrapolating into the future. A worthy goal is to infer physical laws from any correlations you may discover. If indeed science can express the law mathematically, applying that formula should always give a consistently accurate result. Here's my caution:



Drawing conclusions from metadata analysis in the short term—whether about individuals or current events—is both risky and wrong-headed.

In the case of climate change, human activity has been generating more greenhouse gases for only a very short period in comparison with geologic time spans. All we can do today is extrapolate from short-term experience. It might well take several more centuries of data gathering to develop a sample large enough for accurate analysis. However, if the hypotheses about adverse effects were to prove true, by that late date, there might not be time for humans to take sufficient corrective actions.

Objectives of Metadata Analysis

If two data series show similar patterns, a correlation may exist. If there is a correlation, one data series may be the cause, or the generator of change, of the effects on the other. Conversely, if there appears to be no correlation, you must look for a different data series as the cause.

Metadata Analysis and the Stock Market

Much the same process of metadata analysis appears to hold for the U.S. stock market, as represented in the Dow-Jones Industrial Average (DJIA). As *buy-and-hold* investors trust, the long-term trend has been and is consistently upward. That trend is modulated by shorter-term cycles, which may be decades long. Those cycles appear at times to exhibit flat or even bearish performance. Shorter-term variations—on the order of days, weeks, and months—are typical of the volatility that results from events such as news releases, company earnings reports, and changes in federal policy.

As with the other examples in this chapter, making long-term decisions based on short-term results can be both risky and potentially inaccurate. Even a major shock to the system, such as a natural disaster or political upheaval, is not likely to affect the long-term picture. However, in the case of the stock market, cycles and volatility can become decisive considering the comparatively short span of an individual investor's lifetime. Furthermore, the lifespan of the stock market itself has been short (about 200 years so far), compared to other, related long-term trends such as human population growth.

Indeed, the primary limit to DJIA growth probably will be world population. Contributing causes, such as limitations of resources, will control the number of people on the planet. The proviso is *probability of occurrence relies on historical conditions remaining fundamentally the same.* The long-term growth of the stock market has, therefore, a high probability of occurrence. But it is not an assured prediction.

If after careful analysis, if you find the cause—a data series that drives the results of the series under study—you have a basis for a hypothesis. From that tentative rule, you should be able to extrapolate the trend and calculate the probability of occurrences in the future. If you find a cause that always holds true, you may have the basis for a physical law that predicts what will always occur. If, instead, the cause works only some of the time, finding the cause of the cause—the underlying influence of some other variable—may lead to more consistent predictions based on a newly formulated principle or law.



Metadata analysis, in itself, does not yield definitive answers. A correlation and its plausible results provide only a probability of occurrence, not an ironclad prediction. The probability of occurrence will be higher if the conditions present during the time of the study do not change.

Big Data Technology

Big Data technology is a fusion of both new and old repositories of computer-based data.

Internet Technology

The new world is the Internet's World Wide Web, which was designed to be universally accessible and open. Web pages are built using Hypertext Markup Language (HTML), which is text-based. Search engines such as Google and Bing index those pages continually, and search results are cataloged by the most frequently used terms, or keywords.

Text-based searching of large-scale information systems is slow. Search engines overcome this problem by 1) indexing website keywords through background processing, revisiting the pages that change most often and 2) storing hits for commonly searched keywords and key phrases as *binary search trees*. The tree is a stored path from the keyword in the index to the target page. The path is binary because the decision at each branch of the tree is either a 0 or a 1. For example, the route for keyword *chart* might go to *mywebsite.com* if the choice at a branch was 0 or *yourwebsite.com* if it were a 1. Because of the size and complexity of the Web, the binary tree to a target is likely to be a very long string. But because the route is marked at each branch by only two choices, the engine can normally traverse a path anywhere on the Web in less than a billionth of a second. So, when you enter a keyword or key phrase as search text, the engine retrieves the binary tree associated with that key in its index, then executes the path. It does the same for a collection of paths it stores as being most relevant for the key. It presents the top results a page of links it displays for you.

Enterprise Database Systems

By contrast to the Web, the old world is proprietary and deliberately designed to be closed. For decades since the advent of mainframe computers, large organizations, such as multinational corporations and government agencies, have built huge electronic storehouses. These repositories contain data records of individual customers and

vendors, along with daily transactions among those entities. Also on file are reports generated from those records, as well as strategic information such as contracts, product designs, marketing plans, and legal documentation.

In the early days of computer technology, database developers and managers did not anticipate today's requirements for widespread data sharing. Legacy database vendors provided comprehensive, proprietary systems that typically included turnkey service for hardware, software, program development, and maintenance. IBM dominated this marketplace for years, then competitors such as Oracle, SAP, and Microsoft emerged and became giants themselves.

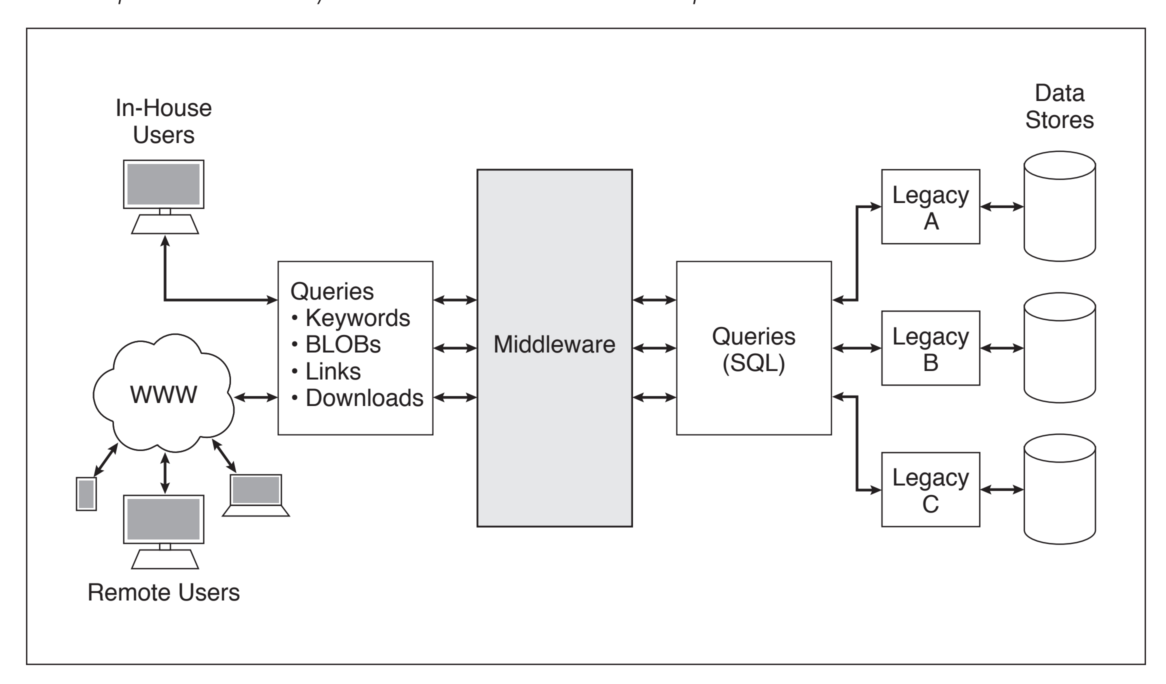
For data custody, the primary concerns of top managers were safeguarding the privacy and security of transactions, along with ensuring competitive advantage. Sharing information with the public was a carefully controlled function administered by marketing and legal departments. And sharing with competitors (or rival agencies) was done with the utmost care, to maintain strategic advantage. Examples of such selective information releases might be advertisements, press releases, engineering articles for trade magazines or professional societies, or required regulatory submissions to government.

Challenges to Collecting Metadata from Legacy Systems

Until the emergence of the Internet in the early 1990s, the concept of global metadata—as well as the scale of technology required to generate it—simply did not exist. But today's large organizations now control planet-wide data repositories and the processing power to manage them. It has become possible to build knowledge bases derived from the enormous amount of information gathered by computer systems over several generations.

Simply converting legacy databases to some universal or more open standard may be neither practical nor desirable. Reasons for maintaining security, privacy, and confidentiality remain. Not surprisingly, administrators in business and government are just as concerned as private citizens about the risks of collecting and sharing metadata. Also, it's understandable that a large commercial enterprise

Figure 16.2 *In this notional diagram of a corporate computer network, middleware handles queries to different legacy systems. These systems house the firm's enterprise databases. The middleware layer also makes it possible for those systems to serve the Web as backend processors for online transactions.*



might want to hoard its hard-won customer experience, yet covet information from its competitors representing industry best practices.

The potential benefits of metadata analysis are considerable. To achieve the integration and ease of data sharing required, a large market niche in enterprise database systems emerged. Appropriately enough, this capability is termed *middleware* (Figure 16.2). Years ago, middleware systems were developed to serve organizations that housed multiple, heterogeneous databases. Such situations arose, for example, when mergers and acquisitions required respective IT departments to consolidate. There was a need to make those systems talk to each other. To promote data interchange, IBM developed a toolset called Structured Query Language (SQL). These days, all enterprise database systems can exchange data via SQL, even though it might not be their native format. (SQL is not the only means of exchanging data among legacy systems, but it is still by far the most common one.)

About the same time as SQL appeared, managers began to complain that their legacy systems were limited to producing predefined reports. They pressed for on-demand information retrieval. But besides overcoming the problem of pulling data from multiple, proprietary systems, there was another major obstacle: Extracting data on-demand from a transaction-based system, such as retail sales, threatened to interfere with those transactions. The solution was the *data warehouse.* Here was a separate middleware system which extracted data into its own repository from legacy systems regularly, typically each night. The executive clients could then play with the information to their hearts' content without the risk of disrupting transactions.

Older middleware technology was slow because of time lags between data extractions. But in today's world, processing capability has grown so large and so fast that middleware is no longer a "snapshot." Middleware systems can provide results in real-time with transactions, or seemingly so. This capability also makes it possible to link legacy systems, or *backends,* with Web-based transaction systems. For example, when you book your airline flight on a website, not only is the airline's legacy reservation system updated immediately so that staff can serve you at the airport, but also the seating chart presented to the next customer is posted to the website so that person can't grab your seat.



I present the above explanation of middleware to show how enterprise database systems have become capable of generating and sharing metadata. What's more, the controls within both the legacy database and in middleware should be designed to ensure that the confidentiality of transaction-level data is maintained. For example, in a middleware system set up as a shared knowledge base, once the summary data is calculated, the transaction data could be discarded. Thus, the middleware system need retain only metadata.

Example: Electronic Health Records

A good example of data collection for metadata analysis is Electronic Health Records (EHRs). Beginning about 20 years ago, hospitals and medical practitioners began to adopt computer-based medical records systems routinely. Until that time, and even today in some offices, patient records were maintained in paper files. An advantage of this approach was a high degree of privacy and confidentiality. Access was

limited to people who could literally get their hands on a file folder. (Fax machines made a difference, but even then the handwritten data wasn't machine-readable.) However, an overwhelming disadvantage was that widespread collection and sharing of medical metadata is difficult if not impossible without machine-readable digital interchange.

One reason for movement to EHR technology is to improve patient care by making it quicker and easier for practitioners to exchange information. Another potent reason is to promote metadata analysis of medical histories and practices. Government healthcare agencies, university researchers, insurers, and pharmaceutical companies pressed for universal adoption of computer-based medical records. For example, with the knowledge bases thus generated, metadata analysis could disclose which treatment plans are most effective, as well as which are most cost-effective. Epidemics and pandemics might be detected as metadata trends. Control measures could be applied before widespread outbreaks.

EHR Risks

In the United States, in the early days of EHR implementation, the Food and Drug Administration (FDA) identified some of the inherent risks, which the agency termed *unintended consequences:*

- Errors of commission (EOC)
- Errors of omission or transmission (EOT)
- Incompatibility between multi-vendor software applications or systems (ISMA)
- Errors in data analysis (EDA)

Of these, EOC and EOT would seem to be the same types of errors that might occur in paper-based systems. An EOC error might result from entry of an incorrect diagnosis code on an input form. An EOT error might occur when a treatment authorization is not received, or not received on time.

Many of the ISMA errors stem from the incompatibilities of legacy enterprise systems. Now those systems must exchange EHRs using

mandated data interchange formats. But those records must then be converted and processed according to the native formats of the receiving systems. ISMA challenges are compounded by the reality that many organization-wide transitions to EHR technology are still underway and not fully tested.

The frequency of EHR data interchange and its sheer volume no doubt contribute to the likelihood of error. However, it's in EDA that metadata analysis would seem to increase risk.

Of the FDA warnings, EDA errors may occur for the same reasons metadata analysis can be misapplied.



Analysts and practitioners may try to generalize trends based on components that are either volatile or cyclic and not causative over the long term. Such a mistake is not so much a Liar's Trick as an unintended consequence, or inadvertent error. In a research environment, the decision may be tempting because there is an unacknowledged bias to produce a particular result.

Benefits of Big Data

Big Data is no doubt here to stay, and it's certain there will be unintended consequences. This topic deserves a book unto itself. No doubt some have already been written and are sure to be revised!

Here are some rapidly evolving technologies that deserve your careful attention.

Searching on Binary Objects

Images, audio clips, and video recordings are a few examples of binary large objects (BLOBs). Rather than searching for alphanumeric strings (text), a search engine can be programmed to look for patterns of zeros and ones. For example, Web apps already exist by which musicians can search for unauthorized postings of their compositions online. Metadata analysis might collect BLOB searches on geologic formations, terrestrial and star maps, DNA, blood chemistry, fingerprints, and human faces.

Image searches, such as facial recognition, are practical today provided the universe of hit possibilities is sufficiently narrow—that is, if the possible targets are limited to pictures of people. But if, for example, you were to search for an image of a candlestick holder, you might instead retrieve a flower vase, a lathe-turned chair leg, or the silhouette of a curvaceous woman.

Dynamic Distributed Networks

Automated Teller Machines (ATMs), cell phone networks, real-time highway traffic reporting, and driverless car navigation are examples of interconnected computer networks in which the configuration of processors, users, and data stores may be continually changing. The much-touted Internet of Things (IoT) promises to be a network of such dynamic networks. The extent to which the IoT will affect our daily lives a wide-open question.

Large-Scale Integrated Manufacturing

Efforts are underway in research organizations to develop an IoT of engineering design workstations, automated machine tools, parts and materials vendors, and robotic manufacturing plants. In such a system, the capabilities of the specific machine tool that will make the part will constrain its design. The machine can be located anywhere on the network. When the part design is approved, the machine tool will be ordered to make it. Accompanying routing instructions will cause the part to be shipped to the plant where robots will assemble the product from a myriad of similarly designed components sourced from all over the world.

Large-Scale Integrated Modeling

The approach just described for manufacturing can be extended to "soft" systems such as agriculture and land-use planning. More comprehensive, predictive models—based on metadata analysis of historical and future trends—will inform the full range of human decision making—from government policy to what crops to plant on your farm next year.

The Potential for Abuse of Big Data

Most of the examples of misuse and abuse of data described in this book arise from unintended consequences. Deliberate fraud, by comparison, is easier to detect. In those cases, there are motive, means, and opportunity. But in cases of neglect or errors of judgment or omission, mistakes can just slip by. And, as I've pointed out repeatedly, it's the commonsensical notions, stemming from assumptions we too often take for granted, that cloud our detective vision.

The unintended consequences of metadata analysis and Big Data are, at this point in the evolution of technology, what engineers are fond of calling *unk-unks*—unknown unknowns. With any tool, be it a wood fire or atomic energy, its sheer power invites both productive and destructive uses. Both will occur, despite control efforts, and therefore invite closer scrutiny and regulation.

It's risky to make generalizations about the average citizen, but I'd venture to say that most people are unaware of the pitfalls described in this book. (The fudging of golf scores is a possible exception!) Years ago, I used to spend hours balancing my checkbook to make sure that the bank had not somehow miscalculated. Does anyone do that anymore? Now there are fewer error-prone manual steps involved in those transactions. The likelihood of arithmetic error is remote. The main risk is failing to spot an unauthorized transaction. But how many people don't bother to question the worksheets output from Excel? If you build a sheet from scratch, do you do test transactions to make sure the logic of its formulas is valid? Or, if you use a template with ready-made formulas, do you simply trust its results because a computer did the work?

The promise of Big Data will be to leverage metadata collection and analysis to inform decisions by artificial intelligence (AI). Doing so may lead to more insightful, well-considered, and effective policies in an astonishingly short time.

If we are so quick to trust PowerPoint, what will we do when this day comes?

> Mr. President: *Computer, what shall we do about the crisis in the Bogus Republic?*

Computer: *Press the button and nuke them all.*

Mr. President: *Why?*

Computer: *You wouldn't understand.*

Okay, this is an extreme example. But what if the computer's recommendation is to sell your business? Discontinue your most popular product line? Declare bankruptcy? If you don't grasp the complexities of the AI solution, will you follow its recommendations anyway?

A Mundane Example

Metadata analysis is the latest iteration in an ongoing trend to assess performance in terms of metrics. "She's doing a good job" has given way to "She exceeded her quota and sold 134 widgets last quarter." There's nothing inherently wrong with this approach. However, managers who ignore subjective factors, or soft facts, do so at their peril. Trusting metrics, by themselves, invites risk. Consider this parable:

The manager obsessed with metrics complained to his supervisor, "Fariba's work output declined by 37 percent in the last 6 months, and her attendance record has fallen below 88 percent." The supervisor replied, "Well, if you walk into the next cube, you'll see that she's 8 months pregnant."

In this hypothetical example (no one I know did this), the manager is not only being obtuse but also insensitive. Depending on the actual circumstances, his oversight might even be grounds for workplace discrimination!

Even with the help of the most powerful AI ever devised, nothing will replace reality-checks from face-time with peers, customer intimacy, and gut feelings.



No matter what the automated systems tell you, don't ignore the "truth on the ground" you get from in-person hands-on experience.

17

How Not to Get Cheated On

Or, What You Need to Know Before the Quiz

THIS **concluding chapter** recaps the clues that a chart maker is cheating. Watch for these tricks, and you won't be fooled! As I've said before, what you do with this knowledge should be a matter of personal conscience. (Remember that they allow you one phone call, and my number is unlisted.)



If you want to be like Hamlet's crab and go through life backwards, you could productively start reading this book here, delving into the previous material as you find potentially useful topics.

Chapter 1: The Numbers Don't Lie—*Do They?*

Labeling can make the difference between having raw data and having useful information. Leaving a label off an important number is a favorite trick of obfuscators and graphic deceivers. Worse, using a label that is only slightly incorrect is the mark of a truly skilled liar.

19 WHAT?

Getting to the infamous "bottom line" is one type of summary, or data reduction. Summaries can make raw data more digestible, but you lose all the flavors of the underlying detail.

Like other kinds of data reduction, arithmetic averages sacrifice detail in favor of a single, more understandable number. And just as relying solely on the bottom line is unwise, basing conclusions exclusively on an average can lead you to incorrect conclusions.

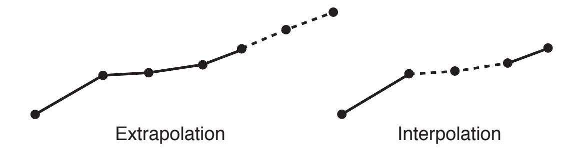
$$\frac{79 + 81 + 78 + 80 + 76 \text{ STROKES}}{5 \text{ GAMES}} = 78.8$$

AVERAGE STROKES PER GAME

There are several other techniques for generalizing about a set of data, all of which might seem much the same to the amateur chart maker and to unwary audiences. Liars can exploit your lack of discrimination by using these terms interchangeably: mean, midpoint, median, and moving average.

These other forms of data reduction have been developed because averages don't always represent typical values. A true average will be more reliable if: 1) there is a large number of individual data values, thus minimizing the effect of any one of them on the result, and 2) the data values are fairly consistent, or without wide variations, thus reducing the effects of unusual values on the result.

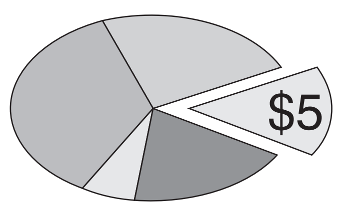
Just as dangerous as data reduction is any attempt to *expand* a set of data, either by guessing in-between values *(interpolation)* or by identifying trends and projecting future values *(extrapolation)*:



The lesson of data reduction is that, as a conscientious chart maker, you can never avoid showing your own interpretation of the data. The trick, if you'd call it that, is to make a chart's interpretation of the data support and reinforce your business message.

Chapter 2: Pies

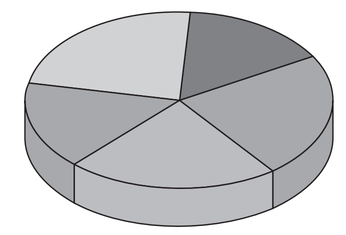
The familiar pie chart is the most overused, misused, and sometimes downright useless trick in the presenter's repertoire. You can use a pie chart to show any proportional relationship between a slice and the whole pie—as a percentage (20%), a fraction (1/5), a ratio (1:5), or a decimal (0.20). The trouble starts when you begin to think of a pie in terms of *amounts*:



Resist the temptation to put values instead of percentages on the slices of a pie chart. If the actual values are essential to your business message, use another type of chart.

Avoid showing two or more pies in the same chart. The almost irresistible temptation will be to size the pies differently, presumably to reflect their different total amounts. This is wrong for two reasons: 1) You will fall immediately into the trap of showing amounts rather than percentages, and 2) Your intuitive notions about how big to make the pies will probably result in exaggerating the importance of the larger ones.

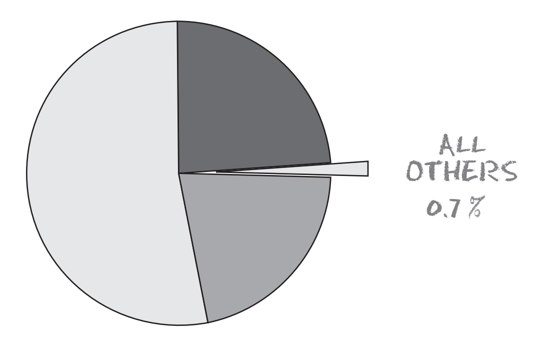
In a 3D pie, people will usually think that the *bottom slice* is the most important. That's because the dimensional effect distorts the apparent size of the slice by literally giving it an edge. People will perceive this slice as bigger than it would appear in a circular, two-dimensional pie:



Liars who want to diminish whatever the pie represents will often clutter the chart with lots of labels, forcing the pie to be smaller to fit on the page or screen.

Liars sometimes find it inconvenient that all pieces of the pie must be accounted for. They typically get around this by labeling a mystery slice All Others. A related trick is to leave a troublesome part of the data out of the pie, so that the whole isn't really whole.

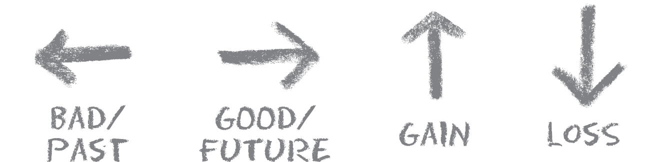
There is one situation in which I recommend that you cheat in preparing a pie chart. If you must include a very small slice—equivalent to 1 percent or less—plot it at about 1.5 percent (about 5 degrees of the circle) so that the thin slice can be seen. Be sure to add a descriptive label nearby that discloses the true percentage:



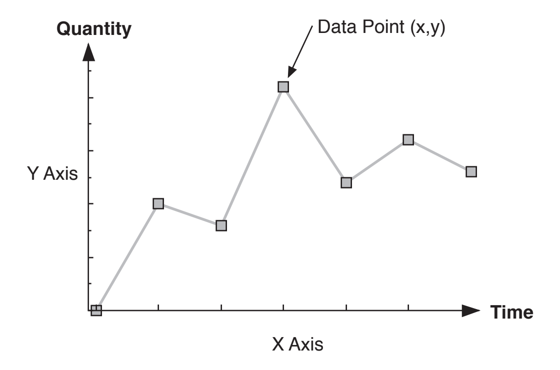
Alternatives to multiple pie charts are donut charts and 100-percent bars. Be cautious about using these, though, if your audience is unfamiliar with these charting conventions. If you intended to show amounts rather than percentages, consider using stacked bars or stacked areas.

Chapter 3: Orientation

Viewers get very different impressions depending on whether shapes appear to go left, right, up, or down. Rightward motion—the way the eye scans a page—is associated with the passage of time—and so with positive movement, and even the idea of progress. Motion upward means gain, and downward means loss.

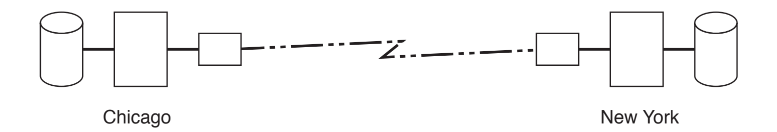


These are the basic assumptions behind the Cartesian coordinate system and the *xy* chart: Time flows from left to right, and the amount of the thing being measured and charted fluctuates up and down.



When your audience shares a set of biases, a presentation will be more effective if its design plays to those assumptions deliberately. Don't try to neutralize their biases—which probably isn't possible anyway. Simply make your designs consistent with what they expect.

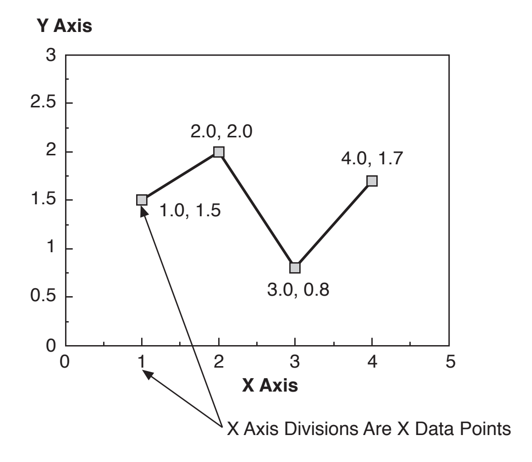
Linguistic and graphic conventions about left-right and up-down carry over to concepts about geography, which we get from relying on two-dimensional maps. Particularly for Westerners, up-down-left-right in the screen can be interpreted as north-south-east-west on the globe. These concepts should also be considered in your graphic designs—including the composition of labels in charts, photographs of plant locations, scenic chart backgrounds, tables of demographic data, and diagrams of computer networks:



Chapter 4: XY Charts

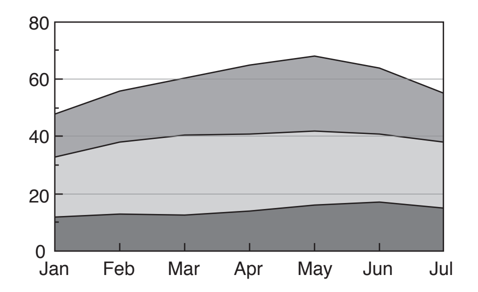
In an *xy* chart that plots quantity versus time, amounts increase or decrease along a vertical scale, or *y axis*. Time progresses from left to right along a horizontal scale, or *x axis*.

In the most common type of *xy* data chart used in business, a sequence of data values is plotted as a series of points on the chart in left-to-right order. The height of each point is its *y* value. Its *x* value is the next division of the *x*-axis scale:



The sequence of points is called a *data series,* or *data set.* Sets of data can be built on one another in two different ways: stacked or cumulative.

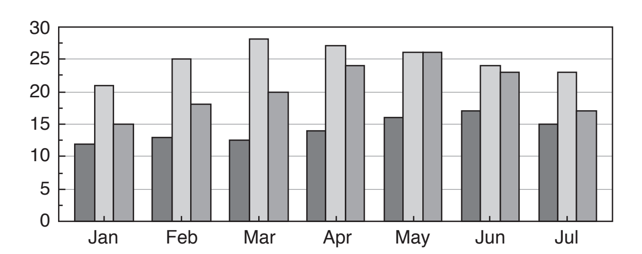
In a stacked area or bar chart, each data set uses the previous set as its baseline. A liar can exploit this characteristic of stacked charts. The fluctuations, or bumps, in the lower data series will enhance and exaggerate the bumps in the upper ones. The honest alternative is to put the series with the least variation on the bottom:



The bumpy effect will be most pronounced in area charts but can also affect perceptions about stacked bar charts.

A liar might also try to stack the lines in a line chart—but this would be a grievous error. The notion of stacking won't be visually apparent, as it is with solid areas or bars. By convention, the baseline of each series in a line chart is usually assumed to be the line y=0 (the x axis).

For bar charts only, an alternative to stacking several data series on the same chart is to cluster them:



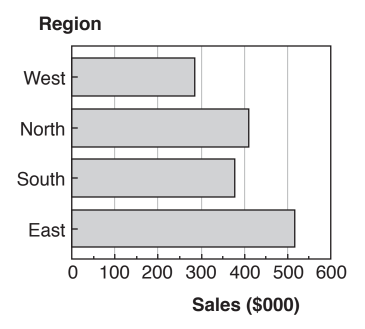
In a cumulative chart, each point within a data series is a cumulative total of all preceding points in the same series. Cumulative charts are a favorite with liars because the general impression is always optimistic: As long as the data points are positive numbers, the trend is always upward—even if some of the increases are small.

Don't confuse *stacked* and *cumulative*. Stacking data series does make them appear to accumulate on top of one another, but this is not the meaning of *cumulative*.

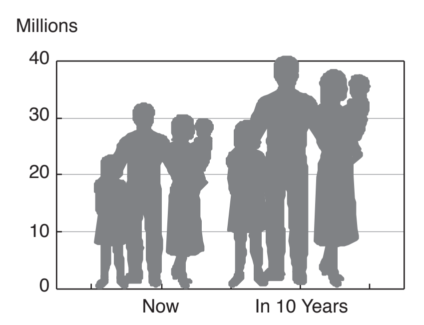
High-Low-Close-Open (HLCO) charts, a special *xy* format for showing stock prices, can be distorted by playing with the appearance of the bars (skinny or fat) and by manipulating the *x* and *y* axes.

In the Gantt charts used for project management, durations of work activities are plotted as horizontal bars. Project managers abuse Gantt charts this way: If the *x* axis is elongated, the audience will get the impression that the overall time span of project tasks has been lengthened.

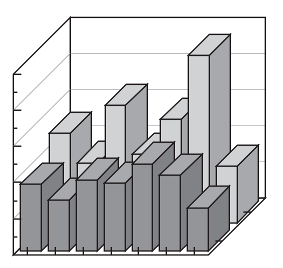
Two of the liars' favorite tricks with *xy* charts are: First, turn vertical charts into horizontal format, subverting the intuitions of the audience about quantity (up-down) and time (left-right):



For your second trick, mix proportional and quantity-time relationships in the same chart. For example, use symbols instead of bars, which distorts their relative sizes:



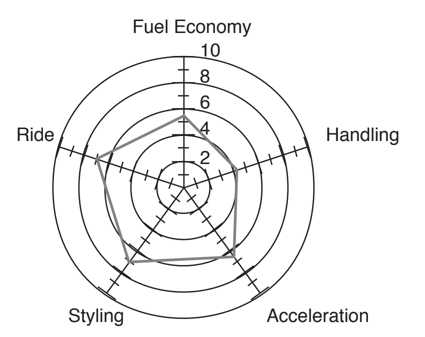
A favorite trick of liars is to use 3D effects to make all the bars in the chart appear taller. The distortion will be most pronounced by a trick of perspective that places the vanishing point above the tops of the bars:



Even if the *x* axis does not represent the flow of time, viewers will regard the bars on the right as the most recent or important. Even if the *x* axis has nothing to do with time, the audience won't be able to avoid associating left-to-right movement with progress. Liars therefore will put the set of bars they want to emphasize on the right end of the chart (but even honest people can do this trick with a clear conscience).

Chapter 5: Radar Charts

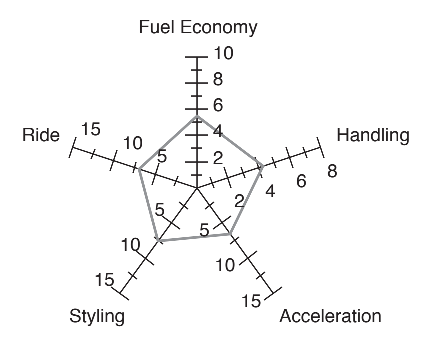
In a radar chart, several axes radiate outward from a common center like the spokes of a wheel. The optional gridlines are shown as concentric circles, each one a uniform distance from the center. A data point is plotted at a measured distance from the center on a gridline. The data point is typically a score or rating for the criterion or category represented by the axis, or spoke.



The radar-chart format imposes a bias on the interpretation of the data: The data set with the most nearly regular shape will be the most well-balanced among the various criteria. Symmetrical plots are therefore "good," and asymmetrical "bad." This bias is just as strong—although decidedly different from—the up-down, left-right biases inherent in *xy* charts.

If you want to be able to draw conclusions from the shapes of the data plots in a radar chart, the scoring must be consistent for all criteria. Furthermore, you should convert all sets of scores to the number of decimal places in the least precise scale. Cheaters might

pretend that the dissimilar scales on their radar charts are proportional to each other:



A particularly nasty trick with radar charts would be to intermix valid charts with questionable ones as adjacent pages or slides in the same presentation.

Chapter 6: Axes

On an *xy* chart, an axis is one of two numeric scales on which the chart maker plots data points. Axes also help readers of a chart estimate the numeric values of its plotted lines, bars, or areas.

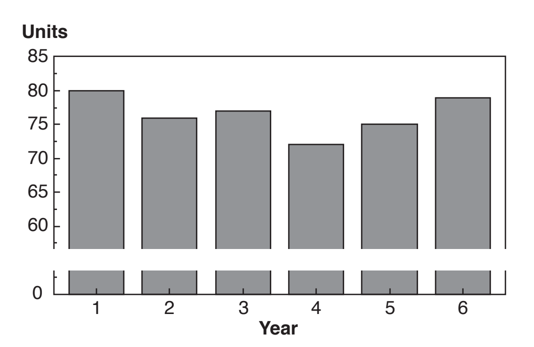
To flatten a plot—minimizing its fluctuations—*increase* the *y*-axis range. To emphasize a plot—maximizing its fluctuations—*decrease* the *y*-axis range. The effect can be heightened if the high point of the plot extends beyond the maximum value on the scale, which unnecessarily exaggerates the magnitude of that data point.



The most truthful version of any chart will be the one that has the least deliberate distortion.

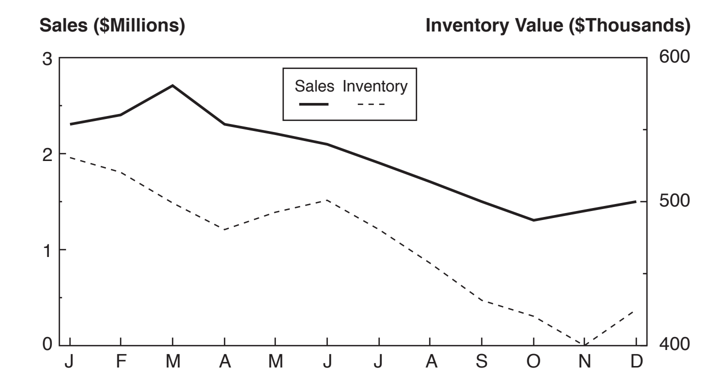
If it suits your lying purposes, don't label the starting point of the *y*-axis scale. Or, don't label any of the scale divisions except the top one. The wider the range and the more divisions it has, the easier it will be to trick the audience into thinking that the scale starts at zero. The

honest alternative is to begin the scale at zero but indicate a break in the scale:

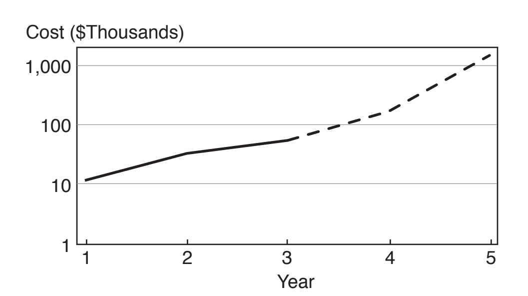


Scaling or using a multiplier on scale values can be a tool for deceitful manipulation of axes. Also, the magnitude of the data can also be either emphasized or minimized by increasing or decreasing the number of scale divisions and their corresponding labels.

Plots of different magnitudes can be shown on the same chart by using two y-axes, with one scale on the left and a different scale on the right. Liars know that such dual-*y* charts can be difficult to interpret if more than two lines or sets of bars are shown—one for each *y* axis. The damage can be compounded if there is no legend, or at least some obvious coding scheme for matching the plots to their respective axes. A clever liar might also use dual-*y* plotting to make bogus comparisons between performance data and some unrelated trend line:



Logarithmic axes, or log axes, are scaled by powers of 10 (orders of magnitude). The legitimate purpose of log scales is to fit a very wide-ranging data set on a single, compact chart. An exponential trend can be made to look less dramatic by plotting it on a semi-log chart, where it resembles a linear trend:



Chapter 7: Choosing Chart Types and Styles

Liars will choose chart types and styles that hide or minimize messages that might otherwise be obvious in the data. That is, liars will deliberately pick the wrong tool for the job.

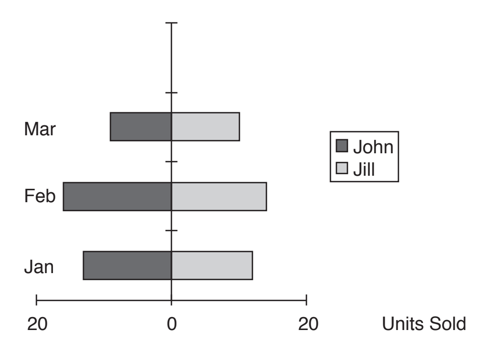
A liar might show you a pie chart, but talk instead about the specific amounts of the total pie or of the slices, drawing the wrong conclusions. In a proportional format such as pie charts, the presenter should stick to percentages and avoid discussing amounts.

Another pie-chart liar's trick is to omit miscellaneous slices, which isn't immediately apparent because the audience will assume that the pie represents 100 percent of the data. Omitting inconvenient slices will also have the effect of increasing the apparent size of all the remaining slices.

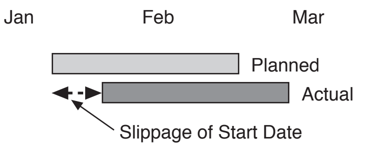
The liar's trick with areas is a fairly simple deceit: Areas simply look more substantial than line plots. Because an area emphasizes total volume, the audience might be less concerned with the peaks and valleys than if the data was shown simply as a line.

Liars also tend to mix stacked and cumulative techniques in the same chart—without making mention of the cumulative part.

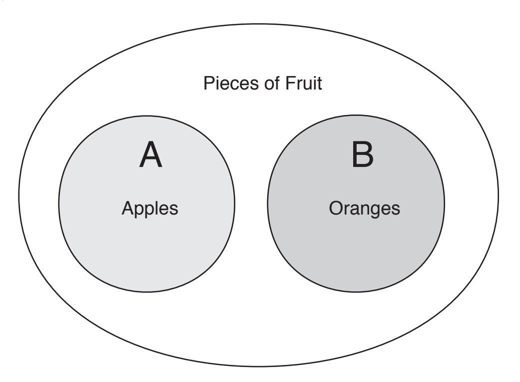
Another liar's trick is to show things that depend on magnitudes—such as sales—as horizontal bars. As a liar might use it, the paired-bar format is an especially confusing use of horizontal orientation:



Save those lazy bars for showing time spans, such as the duration of tasks in a Gantt chart:



Venn diagrams are for showing concepts, not quantities. The sizes of the circles can relate to the relative sizes of the groups, but more often they must be sized to permit larger circles to contain smaller circles, showing relationships between subsets and the larger sets that contain them:



Chapter 8: Trends

Line charts can highlight trends in data better than any other chart type. They are used so often to show trends that audiences are prone to view all line plots this way.

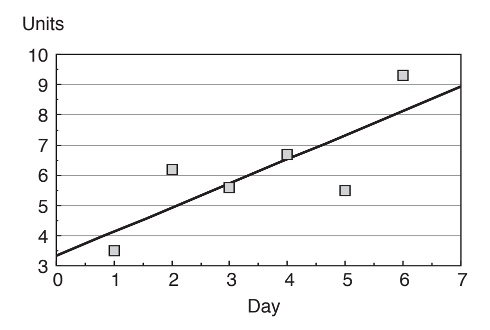
Mathematical formulas can be used to describe trends. Formulas for extrapolation or interpolation work well if the underlying formulas accurately describe the real forces that are shaping the events we observe—especially if the process being described is fairly simple.



Recall from Chapter 1 that liars tend to confuse the important distinctions between *mean, midpoint, median,* and *moving average.*

Error bars can be used to show deviations, or ranges of possible data values. Liars are not usually honest enough to even think of using error bars, much less make the distinction between *population* and *sample* standard deviation.

Regression is another trick of the statistician's trade for finding patterns in data. Common regression types are linear, exponential, logarithmic, and power. By trying different regression formulas, the statistician can fit smooth lines neatly through seemingly erratic sets of data points. The assumption is the better the fit, the more likely that the formula actually describes the process that is generating the data. Regression analysis can be tempting for sincere forecasters and liars alike because it produces smooth line plots and curves that can appear to fit the data neatly.





The pitfall with all kinds of trends is to oversimplify the data.

Chapter 9: Tables and Spreadsheets

How you set up a table in a report or presentation can determine whether it will inform or confuse your audience. Furthermore, if you use a table as a worksheet for chart data, you will find that the arrangement of its columns and rows strongly affects your ability to generate meaningful graphs from the data.

Liars might prefer tables instead of charts. An array of numbers in a table can be difficult to read, and it will be almost impossible for an audience to spot patterns in the data quickly.



As a general rule, you should use a graph instead of a table to present numeric results. Tables are better as worksheets than as formats for presentation.

Even though accounting-style cross-footing is not necessary in electronic spreadsheets, many people still expect to see totals in one of two places—at the right end of rows or at the bottom of columns. So, presenting a table that has an unconventional layout, such as showing the totals at the top, can be disorienting to an audience.

Intermixing data categories in the same column of a table—making it difficult to compare entries in different rows—can be another disorienting trick of the accomplished liar.

In a report narrative or in a speech, you can avoid lengthy explanations of key relationships by summarizing them in a table. If the source data must be shown, consider presenting the numbers in a table beneath the graph:

	Aardvark	Aero	Acme	Able	Archer	Argo	Ace
Last Year	93	71	99	94	121	105	60
This Year	124	83	162	95	143	231	78

Use bullets when presenting topics and key points, especially if you are listing topics that will follow as subtitled sections of your material. Use a numbered list instead if you a presenting a sequence of steps. You can usually increase the size and therefore the legibility of the text in a table if you use fewer columns, perhaps placing explanatory text in footnotes.

Instead of breaking a table, consider placing it on a separate page, even if it must be located farther away from its explanatory text.

Align alphabetic text, such as labels, on the left. Align numeric values on the right, usually with the right-most digit. If numeric values are decimals, align them on the decimal points (or use the same number of places for each string of digits, and right-align all the strings). Optionally, center the headings in their columns:

Company	Index
Acme	253.07
Arrow	4,841.372

The most important consideration in making a chart from a spreadsheet has nothing to do with computer software: You must exercise some careful judgment in selecting which data will be charted. In any sheet, there is almost always too much information to make a single, understandable chart. Out of the whole sheet, you usually need to find just one or two rows of data that tell a story.

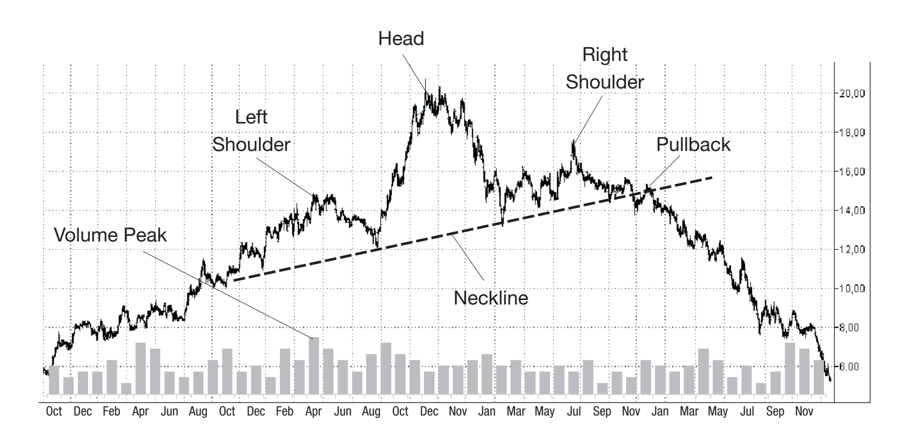
Here are my Eight Great Steps to spreadsheet reliability:

1. Pick the right function for the job: ROUND or TRUNCate?
2. Watch for the kinds of errors the computer can't catch.
3. Make sure to give a function everything it requires (syntax and parameters).
4. Verify all references to data (addresses and range names).
5. Double-check your math.
6. Ask: Do these results *look* reasonable?
7. If you make changes to the sheet, recheck and retest it.
8. If you think your computer goofed, think again!

Chapter 10: Technical Analysis Basics

Technical analysis of financial market charts differs markedly from fundamental analysis, which focuses on root causes and subjective evaluation of corporate performance. In its purest form, technical analysis is concerned only with the relationships between plot shapes as predictors of future market behavior, not as indicators of business conditions. Mixing technical and fundamental analysis can be a pitfall, particularly if you attempt to find underlying meanings and causes in plot curves.

Prognosticators have identified thousands of chart patterns, but understanding Head and Shoulders will give you the underlying principles of support, resistance, pullback, breakout, and projection.



Technical analysis works best on longer-term trends of weeks or months. Short-term patterns, including the microscopic Flash Crash and "crop circles," are not particularly meaningful to investors – unless they have disastrous results, which may be impossible to predict.



Initial speculation about the Flash Crash of 2010 blamed it on inadvertent keystroke errors in entering transaction orders. Subsequent investigation suggests instead a deliberate attempt to manipulate the market. As in all matters of honesty and clarity in financial graphics, whether the cause of error was intentional or not, the consequences can be equally damaging. If you're diligent about catching your own errors, you might not end up an inadvertent liar!

Chapter 11: Viewports and Dashboards

There are cases when a chart may not be misleading in itself, but the method of display introduces some distortion. Such distortion can be particularly troublesome when multiple charts are shown on a page or screen, inviting comparisons among them. Charts are usually scaled to fit a viewport or dashboard, not the display to the requirements of the chart data. Results can be difficult to interpret or misleading.



Comparing *xy* charts side-by-side can be tricky, unless care is taken to make their separate scales proportional to each other.

Chapter 12: Words for Your Charts

Inadvertently leaving off helpful labels on a chart is the quickest way to perjure yourself. Categories of things can be indicated by labels, which are as necessary to reliable data as the numbers, or values.



As a rule, wherever you see a number, you should find a label nearby.

Labels on data can sometimes be hidden or implied. Certain conventions for displaying results, such as delimiters and implied field names, are shortcuts to labeling data:

CALL (800) 555-1212
TO GET INFORMATION

It's a liar's trick to mislabel data, in effect, just by using delimiting conventions that are either unfamiliar to the audience or incorrect in the context of the presentation:

06.09.2049
MEANS SEPTEMBER 6, 2049, IN EUROPE

In general, you should use data labels to: 1) Identify a specific data point you want to highlight in your presentation, such as an exceptionally high or low value, or an average, 2) Make sure that the audience interprets the correct value on plots that are difficult to read, such as estimating the height of 3D bars in relation to the chart scale, 3) Show the actual value for a small bar segment or pie slice that had to be made bigger just to be visible, 4) In an *xy* scatter chart, identify actual values from which trend lines were calculated (such as by extrapolation, interpolation, or regression), and 5) If the design does not permit locating the label close to the data point, use a callout line to direct the eye.

Here are some guidelines for labeling chart axes: 1) Include both the type of measurement and the unit of measure in an axis title. 2) You can omit an axis title only if it would be obvious from the scale labels. 3) If a scaling factor, or multiplier, is needed to interpret the scale labels properly, it must be included either in the axis title or in a footnote to the chart. 4) If an *xy* chart has dual-*y* axes, you must show an axis title for each. Ideally, the axis titles and legend entries should be color-coded to their respective data series.

Here are some other tips for labeling axes for better readability: 1) Use a skip factor to omit labels and regular intervals from a sequence. For example, if the skip factor were five, every fifth label would be printed on the scale. 2) If after using a skip factor the labels still seem too cluttered, adjust their orientation. The normal orientation of labels is horizontal. Some of the alternatives are: vertical up (not recommended), vertical down, slant up, slant down, and stagger.



It's a liar's trick to use color codes that are too similar in legends, making it difficult to correlate the names of data series with their plots.

Use footnotes in charts and in written reports to identify sources of information. The liar's trick with footnotes is pretty obvious: They don't use them!

Font attributes are ways of varying the appearance of a typeface, which can be done to enhance or clarify the labeling in a chart. Attributes can be applied to individual characters, or letters—or to groups of letters, such as words or even entire paragraphs or documents. Also, attributes

such as bold and italics can be used instead of color in black-and-white presentations to highlight labels and emphasize explanatory text. Examples of font attributes include bold, italic, underline, double underline, strikethrough, color, hidden, small caps, all caps, spacing, superscript, subscript, and drop shadow.

When selecting fonts for charts: 1) Use a proportionally spaced font. 2) Use sans serif fonts for chart labels, and serif fonts for titles and explanatory text. 3) Don't intermix fonts within the body of a chart. 4) Use upper- and lowercase letters.

For clarity, use correct punctuation and grammatical style in labels and explanatory text. In your report narrative, be sensitive to the use of pronouns.



Liars are prone to use the inclusive *we* in their speeches and reports as a way of associating themselves with known experts or reliable authorities.

When you are writing your report: 1) Avoid words having more than three syllables. 2) Keep sentences to less than 20 words. 3) Don't include more than two clauses in most sentences.

Professional speechwriters use title screens called *signposts*, which list topics to reinforce the structure of your presentation. Build-up sequences can be used to avoid the "read-ahead" problem when presenting lists:

Subjects of My Paintings
In Southern France

- Landscapes
- Still life

Subjects of My Paintings
In Southern France

- Landscapes
- Still life
- Live models

If you say any words that appear on the screen, make sure that the syntax, or word order, of the spoken words matches the text on the screen. A liar might not bother to signpost a presentation, or might do the opposite—list topics that are never discussed in detail.

Think of your speech or report as a long caption for your charts. If you say something that doesn't have a picture to go with it, come up with a new picture—or get rid of the material.



Liars know a trick we should all imitate: *Make all your mistakes in a loud, clear voice!*

Chapter 13: Financial Proofreading

The requirements of financial proofreading differ from other kinds of copyediting. Rules not only of style but also of consistency apply. Opinions about style differ, but having made those choices, applying them consistently is mandatory.



Inconsistency = Perceived Sloppiness = Lack of Thoroughness = Mistrust of Your Results!

If you're consolidating multiple financial statements in a report, beware of mixing fiscal and calendar year-end figures, audited and unaudited results, or partial-year and estimated figures. Any such discrepancies should be disclosed, usually in a footnote.



Liars, whether they intend to or not, regard email communication as informal and tend to be careless about the wording of their messages. Be concise but clear: Treat email as formal documentation that forms an audit trail and potential evidence in court. Sloppiness, vagueness, and brevity in informal language can create ambiguities. Those mistakes might not be intentional, but might an objective observer think you were being deliberately vague?

Here's your financial proofreading checklist:

- Numeric consistency
- Punctuation and style of English composition

- Appearance, including graphic styles
- Accuracy

You want your reports to give the same impression as your best business dress:

PROFESSIONAL AND CLASSY!

Chapter 14: Color

Biases about color have to do with a person's professional and cultural background. But notions about color aren't that simple: Our interpretation of colors can change with the situation. People make these mental shifts frequently and without conscious effort, usually depending on their understanding of the purpose of a visual message.

Color can be used for 1) a kind of coding, 2) the rendering of real-life pictures and artistic expression, or 3) as a design element of graphic materials, including charts. Business presentations often include all three types of uses.

When colors are used as codes, the color scheme must be as simple as possible, with the least number of colors necessary. A liar might well use too-similar colors in the legend of a chart—attempting, for

Colors Are More Than Just Pretty

To an accountant, green money is good and red ink is bad. Doctors want it just the other way around: oxygen-rich red blood is very good, green necrosis very, very bad! But the situation is also important. The accountants and the doctors will agree on what to do when they see traffic lights on the drive home.

Liars are skilled at sensing the mood of an audience and then manipulating it to their own ends. A keen sense of how colors affect moods is therefore an essential skill for designers of presentations. Yes, perhaps using corporate blue is usually a safe plan. But it's just possible that using a lot of red in making an investment pitch to a group of doctors might stimulate more excitement than alarm.

example, to blur the distinction between otherwise unlabeled plots of profit and loss.

The need for richness and flexibility in color selection when you are showing product photography is just the opposite from the requirement for color codes. Instead of restricting the number of colors, you need to have all the possible subtleties of hue, shading, and tint.

Liars really don't care whether their color presentations are reproduced well in the black-and-white handouts they leave behind. (That's assuming they have the courtesy to give you any handouts at all!) A really nasty trick would be to show a critical plot in light blue—which will disappear completely when printed in grayscale.

When you are designing a business presentation, the range of colors you need will fall somewhere between the extremes of natural color and color as coding. Your preliminary decisions about a color scheme for your charts should include selections for background, main title and subtitles, text (explanatory notes, lists, and chart labels), highlighted text, and subdued text. Background color will be your single, most-important choice.



Liars are fond of using pretty pictures—showing attractive people, flashy vehicles, or exotic vacation resorts—as backgrounds for less stimulating, but perhaps much more significant, material.

Chapter 15: Social Media

Fake news works because it leverages emotion. The immediate visceral reaction from reading a post (or seeing its graphic) confuses any logic you might otherwise apply. Even after you've reasoned away your unreasoned response, the initial emotional impact may persist in your psyche.

Specific Liar's Tricks for concocting posts and confabulate or confuse include:

1. **Change the order of events.** Doing so may reverse inference of cause and effect.

2. **Accuse the accuser.** Blame your opponent before the inevitable blame points to you.
3. **Stretch the analogy.** Point to similarities that don't exist between two stories to support your argument.
4. **Focus on a side issue.** Let the shiny bauble in your left hand draw attention away from the blood on your right hand. Use fiery emotion to fuel a different fire.
5. **Release late Friday (or after the closing bell).** Give bad news when people are likely to be preoccupied with other things.
6. **Confuse with "alternative facts."** Ignore the distinction that differing interpretations of the facts are not facts but opinions.
7. **Point to anomalies as trends.** Make a big deal about a little blip in the curve (and review the information in Chapter 16 on metadata analysis).
8. **Use a pitchman.** Rely on the truism that a song and dance will always be more interesting than a chart.
9. **Animate your charts.** Shorten viewing time and attention spans so no one has a chance to give careful study to your graphic.
10. **Use nonstandard chart formats.** Be wildly creative! No one will understand what you mean.
11. **Abuse the Net Promoter Score.** Capitalize on the dirty secret that asking customers what they think, in itself, may make them think you care and raise the value of your brand.

Chapter 16: Metadata and Big Data

Metadata analysis seeks to find correlations in the curves of two or more data series. The data should be large-scale, long-term, and time-dependent. A long-term plot will typically comprise three types of variations: 1) short-term volatility, 2) recurring cycles, and 3) long-term trend. Identifying the long-term trend requires stripping away the volatility and the cycles. If the long-term trend curves of two time

series appear to match, one of the series may be the cause of the effects in the other. If no correlation exists, the data series under study is not causative. As well, it may be necessary to look for the cause of the cause.

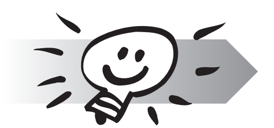
A variable that causes consistent, predictable effects in another time series may be the basis for a new physical law.

Mistakes in metadata analysis can arise from supposing that short-term effects suggest or describe long-term results.

The definitive book about Big Data has yet to be written. We simply don't know enough about its applications or its impacts. But like any powerful, new technology, it's sheer power will tempt abuse.

Go Forth and Multiply...

...And add, subtract, divide—do all the mathematical magic you can conjure. I've spilled my guts, and now you know what I know, which might be just enough to get you into trouble! If you dare to reveal your numbers in a chart, remember that you can no longer pretend to be one of the uninitiated. If your trends are too trendy, your bars bizarre, or your orientation askew, the prosecution will nail you if the jury finds out that you read this book!



I'll say it again: Learning how to lie with charts will help you spot deceptive tricks, whether in someone else's charts or in your own. Where you go from here is up to you!

Unexpurgated Glossary

100-percent bar A proportional plotting style (an alternative to multiple pies) in which data series are shown as stacked bars, each having the same height, and each representing 100 percent of the series total.

A

absolute A value expressed as an explicit amount in some unit of measure; in mathematics, a value expressed as its positive distance from zero.

abstraction In mathematics, a data summary, which trades off detail in favor of an overview or generalization.

additive color mixing The result of adding red, green, and blue to render color tones on a video monitor or in a color film recorder.

address In a spreadsheet, the column letter and row number of a specific *cell,* or data storage location: A1 is the first (upper-left) cell in a sheet.

amount An explicit value of a data item, expressed as a number in some unit of measure, as contrasted with a percentage or ratio of that amount to the total of its data series; quantity.

anomalous Statistics jargon for describing results that defy explanation.

approximation In mathematics, an inexact value; data reduction by summary, rounding, or truncation.

area In geometry, the product of a rectangular object's height and width; a chart type in which the portion beneath an *xy* line plot is shown as a solid object.

argument In the *syntax* of a spreadsheet function, a reference to a data item, such as a cell address or range name, that will be processed by the function.

ascender In typesetting, the portion of a lowercase letter that extends upward from its body, or main portion.

attribute A characteristic, such as color, of a font or graphic object.

audited Describing financial statements that have been reviewed by independent certified public accountants

auto-scaling Process of sizing chart axes in software to fit the range of the data set to be shown

auto-signature In an email application, automatic appending of the sender's signature and contact information to a message

average A kind of data reduction produced by summing a series of numbers, and then

dividing by the count (the number of items in the series).

axis (plural *axes,* pronounced *ak*-seez) In an *xy* or a radar chart, a dimension representing a measurement or criterion according to which one coordinate of the data will be plotted; see also *Cartesian coordinate system, scale.*

B

bar-line An *xy*-chart plotting style which plots some data series as bars, and others as lines.

bar An *xy* chart format in which data points are indicated by the heights of separate graphic columns, or bars.

baseline The reference line in a chart from which distances of the *y* coordinates of data points are measured; in an *xy* chart with no stacking, the line $y=0$; in a stacked chart, the baseline is formed by the plot of the preceding data series; in typesetting, the line on which letters sit.

bearish Describing financial market conditions that suggest declining value over time

Big Data Technology and industry for generating, analyzing, and applying global metadata.

binary search tree Database search path that stores the network route from a keyword to a target page as a string of zeros and ones representing the binary choice at each branch of the tree.

bitmap In computer graphics terminology, an array of colored dots; a picture.

bound Maximum or minimum value on a chart axis or data series.

breakout In technical analysis of market prices, a point at which a stock price departs from a pattern

bubble A chart type for showing the relationship of *sets* against time and complexity or some other pair of *xy* scales.

build-up sequence Division of an on-screen text list into a series of screens so that the list is revealed one topic at a time, as covered in the presenter's speech.

bullish Describing financial market conditions that suggest increasing value over time

buy-and-hold Describing investors who rely on the long-term tendency of a financial market to rise and thus resist timing their trades based on short-term market fluctuations.

C

callout In a chart, a line or arrow that connects a chart object such as a pie slice with its descriptive label.

candlestick price indicator Has four data points: High, Low, Open, and Close.

Cartesian coordinate system In trigonometry and analytic geometry, a scheme of charting (named for French mathematician and philosopher René Descartes) that locates points in physical space according to three perpendicular dimensions, or axes: width (x), height (y), and depth (z).

category A set of criteria by which data can be sorted; in mathematics, the equivalent of a *set*; in a computer database, the equivalent of a *field.*

cell One address in a spreadsheet, or the intersection of a column and a row, which can hold just one data item.

character set The collection of letters, punctuation marks, and symbols in a specific font.

chart A graphic that shows one or more data series as objects; a graphic screen or page in a presentation.

chart type A principal charting format: *xy*, pie, radar, Venn, bubble; see also *plotting style.*

chroma Color purity; see also *tint.*

circular reference In a spreadsheet, any formula that refers to itself, whether directly or indirectly.

clip art Readymade artwork sold for use in graphic materials such as presentations, typically on digital computer media such as CD-ROM.

cluster In a bar chart, a style in which data series are shown as separate bars, grouped at major scale divisions.

CMY or **CMYK** A system of defining color in printed media such as magazines as a subtractive combination of cyan, magenta, yellow, and black; four-color printing process.

color An attribute of the way objects reflect light, an element of visual design; a specific color *tone,* a combination of hue, tint, and shade.

color film recorder A computer output device that creates images on photographic slide film.

color gamut A range of color tones that can be reproduced by a specific device, such as a video monitor or a computer printer.

color palette A range of color tones in a computer graphics program from which a designer can pick tones.

color scheme A designer's selection of color tones for use in a presentation.

color wheel A circular arrangement of *RGB* and *CMY* colors used by graphic designers in making color selections.

column In a pie chart, a second data series shown as a stacked bar, which itemizes one of the pie slices; in a table, all of the positions (cells) under one heading, or field name.

composition In graphic design, the arrangement of graphic objects on a page or screen.

compounding In a financial investment, periodic combination, or reinvestment, of interest with principal so that interest in the succeeding period is calculated on the reinvested total.

constant In mathematics, a fixed amount; not variable; see also *scale factor.*

contrast range A difference in light intensity between the lightest and darkest color tones in a picture.

coordinate One of two or three values (x, y, or possibly z) with which a data point will be plotted in an *xy* chart, based on the Cartesian coordinate system.

correlation In mathematics, the causative link between *dependent* and *independent variables.*

count The number of items in a data series, which can be distorted depending on whether the items include zeros or blanks.

credit line An explanatory note on a chart or photograph that indicates its source.

crop circle Short-term, patterned anomaly in real-time market data feed

cross-foot An accounting technique for verifying the accuracy of a total: adding both down the columns and across the rows of a numeric table should produce the same *grand total.*

CRT A cathode ray tube; a display device in a television or computer monitor.

cumulative A plot in which each point is a running total of all preceding points in the data series; *not* a synonym for *stacked!*

curve In charting, any trend line, whether straight or curved.

custom colors A set of color tones not contained in a predefined color palette; a mixed color.

customer intimacy Marketing term for familiarity not only with customers' product or service requirements but also with their lifestyles.

D

dashboard Display of multiple viewports on a single computer screen for purposes of executive summary

data The plural of the Latin *datum,* meaning *fact;* often used as a singular noun to refer to a collection of data items; a series of data items plotted in a chart or shown in a table.

data item One point or value in a data series; in mathematics, a member of a set.

data processing The means by which raw data is transformed into useful information.

data reduction Using arithmetic to summarize a detailed set of data, with the purpose of making it simpler and easier to understand, often resulting in a single number.

data series or **data set** A collection of numeric measurements plotted on a chart.

data summary In mathematics, a data reduction, or abstraction by which detail data is consolidated for analysis and presentation

data table In a chart, a supplementary display of data values, grouped by columns and rows, often incorporating the chart *legend.*

data type A restriction on a category of data item requiring entries to be alphabetic, numeric, alphanumeric (text, also called *string*), logical (true or false), or in the specific numeric form of an integer, a decimal, and so on.

data warehouse An database that aggregates performance data from multiple systems for purposes of analysis by corporate managers. A data warehouse typically contains overnight or even longer-term results rather than real-time data.

database A specific format for tables in which each column heading is a field name for just one data type, such as a phone number, and each row is a record describing a single entity, such as a person.

delimiter A punctuation or mark that indicates the division between data *fields* in a label or in a data record.

delta (written as the Greek symbol Δ) In mathematics, the difference between two values: $y_1 - y_2 = \Delta y$; see also *slope.*

dependent variable In mathematics, a value that necessarily changes when some other value under consideration changes.

descender In typesetting, the portion of a lowercase letter that extends below the baseline.

diagram A drawing that shows things or concepts as linked symbols; a schematic drawing.

diameter The full height or width of a circle; in geometry, any chord of a circle that passes through its center.

dimension A direction in physical space; in charting, one axis of an *xy* chart.

disinformation Deliberate misinformation; fake news.

dithering On computer displays, the juxtaposition of two dots of different colors to give the impression of a third color.

division One marked unit on the scale of an *xy* chart axis.

donut A charting format that resembles a set of concentric, hollowed-out pie charts.

dual-*y* axes Characteristic of an *xy* chart that has two separate *y*-axis scales, usually left and right, for plotting data series measured in different units on the same chart.

duration A span of time; in a Gantt chart or project schedule, the length of a task or activity.

E

ellipsis In typesetting, three dots (…) used to indicate an omitted word or phrase.

em dash, en dash In typesetting, the long dash (width of the letter *M*) and the short dash (width of the letter *n*) in a font.

entity In data processing, a person, organization, or thing described in a database record.

error bar In *xy* charting, a graphic technique for indicating a range of uncertainty or deviation.

explode To pull one or more slices of a pie chart outward from its center for emphasis.

exponent A power by which a number is raised, as indicated by a superscript: $2^3 = 2 \times 2 \times 2 = 8$.

exponential A data series that increases by a power, or *exponent*; increasing by *orders of magnitude.*

extrapolation Projecting future values from a current series of data.

F

fact checking In report writing, the investigation of cited facts for accuracy by reading reference materials or by contacting witnesses or authorities.

fair use A legal term for quotation or reproduction of small portions of copyrighted material without formal permission from the copyright holder.

fake news Deliberate misinformation; disinformation.

Fibonacci set A sequence of numbers in which each number is the sum of the previous two

field In a *database* table, a heading that describes the data entries in a column; an entry in a data record, or row, of a database.

financial analysis Process of monitoring and evaluating market results, which may employ technical analysis of chart patterns, fundamental analysis of company performance data, and a variety of other methods

fiscal year Year-long accounting period established by a business entity, such as a corporation, for reporting purposes. The start and end date of the fiscal year are effectively arbitrary but must be consistently applied once established. The term is used as distinguished from *calendar year.*

floating bar A bar that does not begin, or have its minimum *y* value, at the chart baseline.

font A typeface in a specific size; generally, a typeface irrespective of size.

foot An accounting term for adding, or summarizing, a column of numbers.

formula A math problem that can be contained in a cell of a computer spreadsheet, referring to values that are held in other cells in the sheet.

four-color process A printing technology that uses the *CYMK* system of color.

function In spreadsheet terminology, a formula that is predefined in the program.

fundamental analysis Financial analysis process which evaluates subjective factors such as quality of leadership, competitive position, and perception of the value of brands

G

Gantt chart A scheduling chart with time spans, or task *durations*, plotted as horizontal bars;
named for industrial engineer Henry Gantt.

gerund The –ing form of a verb used as a noun

grand total In an accountant's ledger, the one total that summarizes all the other totals in a table.

graph An *xy* chart; any chart that shows one or more data series as plots or curves.

grayscale A range of monochrome shades.

H

handicap In golf, a clever technique of scoring that compensates for differences in the skill levels of competing players; a number of strokes by which a golfer is permitted to subtract from his or her score; a shameless liar's trick.

head and shoulders Market data pattern used in technical analysis that shows three peaks, with the middle one highest

high-low-close-open (HLCO) A special *xy* chart format for plotting stock prices.

horizontal orientation An exceptional form of *xy* chart in which the *x* axis is vertical, the *y* horizontal; see also *vertical orientation.*

HSL A system used by graphic artists to describe color tones in terms of hue, saturation, and lightness.

hue One of the primary colors from which color tones are blended.

I

independent In mathematics, a variable that is not necessarily affected if the value of some other variable under consideration changes.

intensity Another term for color *shade.*

international characters Accented letters and symbols contained in a font for composing text in Romance languages.

interpolation Estimating the missing, or in-between, values in a data series.

intersection The values or data items that two *sets* hold in common.

iteration In spreadsheet terminology, the repetition of a calculation (also called *pass* or *loop*).

J

Jones The guy who wrote this book; American slang for an irrepressible urge.

K

key See *legend.*

keyword In spreadsheet terminology, a reserved name used as an argument to request a specific type of result from a *function.*

L

label Text displayed in a chart that names or explains a displayed value or plot; in spreadsheet terminology, any text entry.

legend In a chart, a key that correlates the names of data series to their plots, usually by color codes or patterns.

letterform In typesetting, the design of a character in a font.

letterspacing In word processing and typesetting, the white space between characters or between words

lightness Another term for color *shade*.

line An *xy* chart format in which data points are connected to form a continuous, segmented line (polyline), or *line plot*; a geometric object defined as the shortest distance between two separate points in a plane.

linear Describes a straight line, which is the shortest distance between two points; an *xy* chart that is not logarithmic; see also *logarithmic*.

linked bars An *xy*-chart plotting style in which tops of bars are connected by line segments to emphasize trends.

list An unstructured collection of rows that contain dissimilar but perhaps related data records, as contrasted with the structure of rows in a table or database.

logarithmic Refers to an *xy* chart in which one axis (*semi-log*) or both axes (*log-log*) have been scaled by powers of 10; see also *scaling*.

log-log An *xy* chart in which both axes have been scaled by powers of 10.

M

major division A tick on a chart axis scale that marks the primary increment of its unit of measure; in all types of *xy* charts except scatter charts, the ticks on the *x* axis at which data items are plotted in sequence—from left to right in vertical charts, from top to bottom in horizontal charts.

mean See *average*.

Mean Time Between Failures (MTBF) Average estimated time between required repairs, based on product-failure history.

median In a data series, the item located exactly in the middle of the count of items, where there are the same number of values above it as are below it; if there are an even number of items in the series, the median is the average of the middle two items.

metadata Data that summarizes other data; large-scale data reduction, typically over lengthy time spans in demographic populations.

metadata analysis Process of finding long-term correlations between metadata series by removing volatility and cycles to disclose trend curves.

metric Synonym for any kind of measurement

midpoint The average of just the high and low values in a data series.

minor division A tick on a chart axis scale that marks the secondary increment of its unit of measure; the subdivision of a major tick interval on a chart scale.

misinformation In contrast to disinformation, which is deliberate, incorrect information due to inadvertent error or misjudgment.

monospace In typesetting, a font in which the spacing between letters is uniform.

mood The emotional response of an audience, which can be affected by on-screen colors.

moving average In an ongoing series of values, the arithmetic average of only a specific number of recent ones; for example, in a lifetime of golf scores, the most recent four-game average.

multivariate analysis An investigation of several factors or criteria that contribute to a single result or that describe a single entity (one purpose of radar charting).

N

neckline In technical analysis, a reference line drawn through the troughs, or lowest points, of a head-and-shoulders pattern

nesting In mathematics and in computer programming, refers to a process, such as an equation or formula, that is contained within a larger process; use of sets of parentheses in algebra or in spreadsheet formula syntax to control the order of processing.

Net Promoter Score (NPS) Survey methodology that asks how willing the customer is to recommend the company's products or services to someone else.

O

object In computer graphics, a basic drawing element, such as a string of text or a geometric shape; a collection of graphic objects that can be manipulated as a group.

order of magnitude A power of 10; for an integer or decimal value, the number of digits to the left of the decimal point, minus one—for example, the order of magnitude of 160,000 is 5 (1.6×10^5).

ordinal Describes a data series for which the sequence of items uses some method of counting: 1, 2, 3...or January, February, March...

orientation The position of a chart element in relation to something else; with respect to the chart frame, the *y*-axis options *vertical* or *horizontal;* in a presentation, the way in which a rectangular sheet of printed paper or a display screen is to be viewed—*portrait* (long dimension vertical) or *landscape* (long dimension horizontal); in graphic design, the position and direction of graphic shapes in relation to a page or screen; see also *composition*.

origin In the 2D Cartesian coordinate system, the point at which x=0 and y=0, or (0,0); in a vertical *xy* chart in which both scales start at 0, the lower-left corner of the axes.

outlier Anomalous point in a data series.

overlay A clear sheet used by graphic artists as a separate drawing layer; in a 2½D Cartesian coordinate system, a drawing plane.

P

paired-bar A horizontal *xy* plotting style in which two comparable data series are shown as opposed sets of bars—one plotted from the center rightward, the other from the center leftward.

parallel construction Grammatical principle by which, in a list of items, all elements follow the same style rules

pattern In technical analysis, the distinctive shape of a line plot in market data, which may reveal events that tend to recur and from which projections might be made

perspective A two-dimensional drawing technique used to render the third dimension, or depth, as lines that converge on a distant point on the horizon; see *vanishing point.*

pica A typesetter's unit of spacing, equal to ⅙ inch.

pie A circular or elliptical chart in which percentages of the whole are shown as slices, or sectors.

plagiarism The unauthorized and uncredited use of material from a copyrighted source; see also *substantial similarity*.

planar Having the characteristics of a geometric plane; two-dimensional; flat.

plot In an *xy* chart, the line or curve described by connecting the points in a data series.

plotting style A set of options within a chart type: bar, stacked bar, bar-line; see also *chart type*.

PMS colors Color tones as defined by the Pantone Matching System developed by Pantone, Inc.

point The location of a data value on a chart; on a vertical *xy* chart, a point is defined by a horizontal distance (*x*) and a vertical distance (*y*) from the origin (0,0); the delimiter in a decimal value; a typesetter's unit of measure equal to 1⁄72 inch; see also *scatter*.

POM dependency Sardonic slang term used by computer programmers to describe results that seem to have no apparent, consistent, or controllable cause—therefore, depending on the phase of the moon.

population In statistics, a *standard deviation* or other measurement that includes all items in a data set.

precision In a decimal value, the number of digits to the right of the decimal point.

primary colors In the *RGB* color wheel, pure hues such as red, blue, and yellow.

programmed trading Automated initiation and processing of market transactions based on predefined criteria, typically employed by institutional traders. Programmed trading can respond to changing market factors in a matter of milliseconds.

progress In the graphics of Western cultures, movement from left to right.

projection Extrapolation, as when projecting future stock prices based on technical analysis of market patterns

proportional In mathematics, refers to the ratio of one value to another, describing a percentage; in typesetting, a font that permits variable spacing between letters.

public domain A legal term describing material for which there is no private copyright holder.

pullback In technical analysis, the point at which a stock price returns to the support line, or price floor in the pattern

Q

quantity A numeric value that requires a unit of measure for a label; a desirable feature if you also have *quality;* see also *amount.*

quote stuffing Malicious market trading practice in which a computer program makes a large number of requests for price quotation on a particular stock in a very short period of time in an attempt to run up the price of the stock without actually buying any of it

R

radar A chart format for investigating symmetrical relationships in which axes for plotting multiple criteria radiate from the center like the spokes of a wheel; what they got me with on the road between Phoenix and Payson.

range The extent of a chart axis or data series that lies between its bounds, or minimum and maximum values; a data series; a set of permissible values; in a spreadsheet, a contiguous

block of cells, each of which can contain one data item.

readability index A score that indicates the educational grade or skill level required to read and comprehend a specific document.

reasonableness An accounting term for the appearance of accuracy in results.

record In a *database* table, one row, which must describe just one entity, such as a person. (A row *must* describe one and only one entity.)

regression In mathematics and charting, a technique for fitting a smooth curve through a given set of data points; what will happen to you if you work too hard writing books; technically, a formula in calculus for minimizing the residual square error between data points and the fitted curve in a process called *least squares analysis.*

relative A value not expressed as an explicit number of some unit of measure, but as measured against some other value of the same data type; for example, a ratio.

resistance line In technical analysis, upper bound of a chart pattern, indicating top price

RGB A system used in color computers and video cameras for describing color tones as mixtures of red, green, and blue.

rolling release A series of news releases timed to appear at the optimum time of day in each region or time zone.

rotate To adjust the starting point of a pie chart by some offset from zero degrees, usually 0–360 degrees counterclockwise.

rounding The approximation of a long decimal value to a fewer number of decimal places by dropping all but a given number of digits to the right of the decimal point, possibly adjusting the right-most remaining digit to approximate the discarded digits; if the right-most digit is increased, the process is called *rounding up*, if decreased, *rounding down.*

S

sample In statistics, a *standard deviation* or other measurement that includes a selected portion of all items in a data set.

sans serif In typesetting, a font that has no ornamentation at the ends of its letters, such as Arial and Helvetica.

saturation Another term for color *chroma* or *tint.*

scalable Refers to a font that can be adjusted continuously in size, rather than being restricted to discrete point sizes.

scale A chart axis marked at regular intervals, or *divisions*, according to a unit of measure.

scale break In *xy* charts, a mark on the *y* axis to indicate nonzero starting point

scale factor A constant by which a chart axis scale is multiplied or divided, such as thousands or tenths.

scaling Multiplying or dividing chart axis scale values by a constant, or fixed, amount; see also *scale factor.*

scatter A type of *xy* chart in which *x* values need not be plotted in sequence at major scale divisions, removing any association between the *x* axis and the progress of time; a point chart.

secondary colors In the *RGB* color wheel, hues that are blends of primary colors, such as red-violet.

secondary *y* axis In an *xy* chart, an optional y_2 axis; see *dual*-y *axes.*

segmented bar See *stacked.*

semi-log An *xy* chart in which one axis, usually the *y*, has been scaled by powers of 10.

sense In grammar, the correct interpretation of a written sentence, which should be clear from its composition and syntax.

serial comma In English punctuation, the use of a comma to separate words listed within a sentence.

serial date A decimal form of calendar notation, including the date and time, used in computer processing; the digits to the left of the decimal point represent the date, counting from the year 1900 (or 1904 on some systems), and the digits to the right represent the 24-hour time in hours, minutes, and seconds.

serif A font with cute doodads and curlicues at the ends of its precious letterforms; one such doodad or curlicue.

set In mathematics, a group of related things, usually numbers; a data series.

shade The amount of black in a color tone.

shape The outline of a graphic object; see also *curve.*

sheet Jargon for *spreadsheet* or *worksheet.*

short-sale order In financial trading, a transaction that bets on the decline rather than the appreciation of a stock price.

signposting A speechwriter's technique for listing the major topics of a presentation on the screen.

skew A distortion of the result of a data reduction due to unusually high or low values in the data series.

skip factor A specification for omitting selective labels from the scale of a chart axis; for example, a skip factor of five would cause every fifth scale label to be displayed on the axis.

slice A data item, shown as a percentage, in a pie chart; also called a *cut, piece, segment, wedge* or *sector.*

slip card A piece of cardboard used to hide upcoming presentation topics listed on an overhead transparency.

slippage In a Gantt chart, the visible gap between a task's estimated start date and the actual start date.

slope A slant of a chart line or curve, defined for two data points by the change in y (Δy) divided by the change in x (Δx); the rate of change.

smart quotes In word processing, automatic insertion of curly quotation marks

spoofing In financial trading, malicious practice of entering and withdrawing thousands of orders over a span of hundreds of trading days in ongoing attempts to manipulate futures markets

stacked A charting option for showing data series in bars or areas that are built on top of one another, like a stack of bricks.

standard deviation In statistics, the degree to which a data value differs from the *mean.*

stop-loss order In financial trading, a transaction that triggers a sale at the lowest price the investor is willing to tolerate

style sheet A type of *template* used to create documents; a collection of graphic options used to define the appearance of a document.

subscript In mathematics, a notation used to indicate the sequential number of a data item within a series: x_3; the screenplay for *The Enemy Below.*

substantial similarity The legal term used to define material that has been plagiarized, or used without permission from a copyrighted source.

subtotal In a table, a component of a *grand total.*

subtractive color The result of filtering reflected light through colored dyes on different layers of photographic film or through printing inks overlaid on paper.

subtype Variations within a chart type: area, stacked area, 3D area; see also *plotting style.*

summary A type of data reduction produced by adding a series of numbers; a total.

support line In technical analysis, a reference line indicating the likely minimum price below which the stock won't fall, as long as it stays within the pattern

syntax In spreadsheet terminology, the required written form of a formula or func-tion; in English composition, the structure of a sentence.

T

technical analysis In financial analysis, a discipline by which patterns in market charts can be analyzed to reveal recurring trends

template In computer graphics, a generic design that lacks only the user's data to create a finished chart.

thematic coding The use of color to identify principal topics in a presentation.

three-dimensional (3D) Chart plotting styles such as lines, bars, or areas shown with an edge to suggest thickness or depth.

tick A short line segment; a mark on a chart scale that indicates a major or minor division; a mark on a floating bar in an HLCO chart that indicates the opening or the closing price; in financial trading, the difference between the opening and closing prices on a given day.

time analysis The scheduling process in which the starting and ending dates of tasks are correlated to find the *critical path*, or minimum time span required to complete all essential tasks.

time In a typical *xy* chart, the *x* coordinate, which varies along a left-to-right (horizontal) *x* axis; in a Gantt chart, the span of a horizontal bar.

time series A data series in which the sequence of points is ordered by the flow of time.

tint The amount of white in a color tone.

trend analysis The study of data or plots in charts in an attempt to find meaningful patterns.

trend The identifiable pattern in a series of data items; a chart curve that describes this pattern from which future results can be *extrapolated* and intermediate results can be *interpolated.*

trough In a market chart, low point of a price pattern

truncation The shortening of a long decimal value to fewer decimal places simply by dropping all but a given number of digits to the right of the decimal point; see also *rounding.*

two-and-a-half-dimensional (2½D) Cartesian coordinate system for plotting *xy* charts in which a third axis, *z*, is used to fix the position of multiple drawing planes, or overlay sheets.

typeface or **face** A style of lettering; a font, irrespective of its point size.

typesetting The reproduction of text in print, according to a specific graphic design.

U

unaudited Referring to financial statements that have not been subjected to independent examination

unintended consequences Inadvertent errors, often from trusting the results from automated systems.

union In mathematics and logic, the combination of two *sets*; the marriage of like minds (which admits no impediment).

V

value A number in a data series; another term for color *shade;* see also *label.*

vanishing point In a two-dimensional drawing that renders the third dimension by using perspective, a point on the horizon toward which depth lines converge.

variable In mathematics, a placeholder for an unknown value or any of a range of possible values.

vector graphics In computer graphics applications, an image composed of lines and areas

Venn A type of diagram, named for English logician John Venn, that uses circles or ellipses to show relationships between sets and subsets.

vertical orientation A typical form of *xy* chart in which the *y* axis is vertical, the *x* horizontal; see also *horizontal orientation.*

viewport A chart-display window on a Web page or application screen

visual bias A set of assumptions shared by members of a cultural or professional group about the inherent meanings of graphic compositions and elements.

visual cueing The appearance of text or an object on the screen at the same moment it is spoken or named by the presenter; television newswriters often screw this up because, working against tight deadlines, they have no idea what picture the producer will find to put with their words!

visual priority In 2½D or 3D Cartesian coordinate systems, the *z*-axis value of a graphic object, which determines whether it will cover or be hidden by other objects; objects with low priority (high *z* values) will be in the background and will be hidden where they overlap with objects of high priority (low *z* values), which will be in the foreground.

volatility Referring to the frequency with which a stock is bought or sold during a trading session or over a span of time

volume bars In market charts, a bar chart showing the number of shares traded during the same time period as the price plot shown on the same chart

volume Total amount (volume of financial trading); 3D spatial measurement (height times width times depth).

W

weight In statistics, a factor or multiplier used to compensate for an incompatibility between data sets, such as different units of measure or levels of importance.

workpaper An accounting term for an interim document; scratchpad.

workstation A desktop computer or terminal; technically, a powerful microcomputer used for graphic or engineering applications.

X

***x* axis** In a typical *xy* chart, the horizontal scale against which the first coordinate *x* (usually representing a point in time) is plotted.

xy A chart format based on the 2D Cartesian coordinate system in which time typically progresses along the horizontal *x* axis and measured quantities vary along the vertical *y* axis.

Y

***y* axis** In a typical *xy* chart, the vertical scale against which the second coordinate *y* (quantity) is plotted.

year-end The date on which a company's yearly accounting period ends, which may be calendar or fiscal

Z

***z* axis** In the 3D Cartesian coordinate system for *xy* charts, a third axis that indicates depth; in a 2½D system, the position of the drawing planes, or overlay sheets.

zzzzzzzz A soporiferous condition not likely to be induced by this stimulating discourse on an otherwise dreary motif.

Index

100 percent bar charts, 39–40, 123–124

2D representation in Cartesian coordinate system, 62

2½D representation in Cartesian coordinate system, 62–63

3D area charts, 116–117

3D bar charts
- colors for, 274
- perspective in, 78–80

3D chart formats, 194

3D pie charts, 29–30, 108

3D representation in Cartesian coordinate system, 62–63

A

abbreviations, 218,223–224, 238, 246–247

abstraction, 2, 129, 169, 171, 345

accounting
- color in, 271
- spreadsheet styles, 145–149

accounting periods, 200, 240

accuracy
- in axis scales, 101
- in investment functions, 159–161
- in radar charts, 89–90
- in tables, 157–160

acronyms, 218, 224, 238, 244–245, 247

additive color mixing, 280

addresses of cells, 154

advertising, geographical biases in, 56, 57

air travel, 56–57

alignment
- fonts for, 265
- in tables, 152

all caps font attribute, 211

All Others category, 35–36

alternate spellings, 223–224

American Medical Association (AMA), 233

amounts, 4
- in pie charts, 21–23, 30–32
- in *xy* charts, 47–49, 61, 64–65

ampersand (&), 247–248

analogy, 342

anomalies, 142

apostrophe (’), 215, 245–246, 251

approximations in tables, 159, 161–162

area
  of symbols, 77–78
  in two-pie comparisons, 34

area charts, 66
  as pie chart alternative, 40–41
  stacked, 68, 70, 115–117
  using, 114–118

arguments for functions, 156, 162–163

Arial font, 263–264

arrows, slanting, 53–54

Artificial Intelligence (AI), 317-318

ascenders, 266

Associated Press (AP) Stylebook, 232, 243

assumptions in data reduction, 7

attributes for typefaces, 211, 264

auto-scaling, 190, 345

auto-signature, 253, 345

automated checking, 246

AVERAGE function, 162

averages, 11–12
  blank values in, 14–15
  means, 11
  medians, 13
  moving, 14, 136–137
  in trends, 134–135

axes, 93–94
  bounds for, 95–98
  breaks in, 99
  dual-*y*, 101–103, 113–114
  labels for, 65, 94
  logarithmic scales on, 103–106
  on radar charts, 84, 89–92
  rotating, 75–77
  scale factors for, 100–101
  on *xy* charts, 48–50, 61, 64–65, 75–77
  axis, 48–49, 58, 61–74, 77–79, 170, 172, 189

B

backend, 313

background
  for charts, 278–279
  color for, 266–269, 279
  palettes for, 272–273

bar charts, 9, 10, 65
  clustered, 69, 120, 121
  cumulative, 122–123
  for Gantt charts, 52–53, 73–75
  paired–bar, 124–125
  perspective in, 78–80
  100 percent, 39–40, 123–124
  as pie chart alternative, 39–40
  for sales, 51
  stacked, 9–10, 69–70, 120–121
  for time, 51–52, 77
  vertical, 118–124

bar–line charts, 118–119

bars on HLCO charts, 72

baselines
  for area charts, 40, 41
  for line charts, 68, 70

bearish, 176, 310

bias. *See* cultural bias

bibliographies, 261

Big Data, 301, 310–318

Bigollo, Leonardo Pisano (Fibonacci), 185

binary search tree, 310

bitmap, 192, 279, 346
blank values in averages, 14–15
blocks in spreadsheets, 154–155
blue
  for background, 268–269
  response to, 259
bold attribute, 264–265
bottom line, 8–9
bounds, 91, 94–98, 138, 164, 173–174, 185–186, 190, 346
breakout, 173–174, 176–177, 245, 335, 346
breaks
  in axes, 99, 99
  in tables, 151
bubble charts, 126–128
build-up sequences, 228–229, 338
Bulkowski, Tom, 177, 179
bulleted lists, 151, 242, 244
bullets, color for, 266
bullish, 176–177, 310
bumpy effect in stacked charts, 70, 116
buy-and-hold, 309, 346
buy-or-sell decisions, 174

C

calendar years, 284
callout lines, 36, 202, 211, 337
candlestick, 171–172, 316, 346
capitalization, 221–222
Cartesian coordinate system, 62–63
categories, 3
  All Others, 35–36
  labels for, 4–5
cathode ray tubes (CRTs), xix
cells in tables, 154
changes, emphasizing, 95–99
character sets, 263
characters
  height of, 214
  international, 209
check registers, 146–147
checklist, 239, 339
Chicago Manual of Style, 207, 232, 234, 243, 251–252
chroma, 257–258 , 347
circular references, 160–162
citations, 206–207
climate change, 295, 307–308
clip art, copyright protection for, 208
clock watchers positioning rule, 25–26
clustered bar charts, 69, 120–121
CMY and CMYK color methods, 280–281
coding, color for, 261–262, 264
colon (:), 154, 215, 219, 244–245, 251
color, 255–256
  for background, 266–269, 279
  case study, 283–284
  for coding, 261–262, 264
  with computers, 270–278
  custom, 274–276
  as design element, 264
  designing for, 280–281
  for dual *y*-axes, 102, 257
  with legends, 259, 262

for lists, 282
numbers for, 274
palettes for, 265–266, 272–274, 280
in presentations, 264–266
psychology of, 258–260, 268–269
as rendering tool, 262–263
schemes for, 265–266, 272
on screen, 277–278
selecting, 266–272
for tables, 150
terms for, 256–258
color wheel, 261, 281–283, 347
column charts, pie–column, 37. *See also* bar charts
columns in tables, 143, 145, 148, 151, 154
comma (,), 216, 219–220, 233, 237, 240, 244, 251, 355
company names, 238–239
comparisons
in tables, 149
two–pie, 32–34
compatibility of color output, 269
composition, 43
compound sentences, 215, 219
computers, color with, 270–278
conjunctions, capitalization of, 221
consistency, 223, 231–254, 339
Consolidated Audit Trail (CAT), 183
contrast
with color, 267
in reproductions, 270
control engineers, color response by, 259–260
converting output media, 269–270
coordinates in Cartesian coordinate system, 62
copying table information, 163–164
copyrights, 207–209
corporate style, color in, 256
correlation in metadata analysis, 303, 343
correlation in regression analysis, 140
counting, 2–3
Courier typeface, 212–213
credit lines, 207
criteria on radar charts, 84
crop circles, 184, 335, 347
cross–footing in spreadsheets, 145–146
cultural bias, 59–60
in color, 258–260
considering, 45–47, 58–59
in geography, 54–57
in increases and decreases representations, 44–45
for tables, 144
in time representations, 43–44
in *xy* charts, 47–53
cumulative charts, 71–72, 122–123
custom colors, 274–276
customer intimacy, 300, 318
cyan, response to, 259
cycles, 305, 306, 342

D

dashboard display, 189, 194–195, 336, 348
dashes, 271
data, 3
expanding, 16
vs. information, 6, 282
raw, 2

data points, labels for, 197–205

data processing, 2

data reduction, 2, 7, 301–309
- averages, 11–15
- summaries, 8–10
- trends, 16

data sets, 1, 65, 67

data summary, 171, 345, 348

data warehouse, 194, 312

database tables, 147–149

dates
- calculations with, 162
- commas for, 270
- formats for, 199–200, 216

decimal numbers, precision of, 157–160

decimal points, 152, 158, 241, 334, 352

decimals, 241, 334
- 100 percent charts for, 39–40, 123–124
- pie charts for. *See* pie charts

decreases, representation of, 44–45

delimiters in labels, 198–199, 336, 348, 353

Deming, W. Edwards, 74

demographic, 301–302

dependency, POM, 114

dependent variables in regression analysis, 140

Descartes, René, 62

descenders, 212

designing
- for color output, 280–281
- for monochrome output, 279–280

deviation from mean, 137–139

diagrams, Venn, 125–128

dictionaries, 225, 235

dimensions and axes, 94

direction
- cultural biases in, 54–57
- of increases and decreases, 44–45
- of time and progress, 43–44

disinformation, 285–295, 348–349, 351

distortion, 1, 40, 68–70, 77, 98, 117, 189, 192, 326, 328, 336, 355

dithering, 278

division by zeros, 162

divisions on axes, 65, 94, 100–101

donut charts, 38

double-hyphens, 217

double underlining attribute, 264–265

Dow Jones Industrial Average (DJIA), 172, 181–182, 309

drawings, geographical biases in, 57

dual-*y* axes, 101–103, 113–114, 257

duration of tasks, Gantt charts for, 52–53, 73–75

E

editorial “we,” 224

Electronic Health Records (EHR), 313–315

ellipsis (…), 217

em dash (—), 217, 251, 349

email, 231, 235, 252–253, 287, 297, 339

emphasis
- on fluctuations, 95–99
- italics for, 220–221

en dash (–), 217, 251, 349

Encyclopedia of Chart Patterns, 177
entities in database tables, 148
equal sign (=) for table functions, 156
error bars, 138–139
Excel, xv–xx, 51, 93, 147, 153, 162, 190
executive information systems, 194
expanding data, 16
exploding pie chart slices, 27, 29, 37, 108
exponential regression, 140
exponential trends, logarithmic scales for, 104–105
extrapolation, 16, 132–133, 169, 202, 309

F

faces (human), 60
faces (typefaces), 208–211, 256, 337, 349, 356
fact–checking, 226, 289
fair-use doctrine, 207, 349
fake news, 289–295, 341–342
FALSE table values, 163
Fibonacci series, 185–186, 349
fields and field names in database tables, 148, 198–199, 336, 346–349
financial analysis, 129, 167, 349, 350, 356
Financial Industry Regulatory Authority (FINRA), 183
financial managers, color response by, 258–259
financial market, 167, 170–175, 184–187, 335, 346
financial proofreading, 231–254, 339
financial reporting, 168, 237, 240, 255
Fiscal Year End (FYE), 237
fiscal years, 237, 349
Flash Crash of 2010, 179, 181–184
flat charts, 62
floating bars in HLCO charts, 72
fluctuations, emphasizing, 95–99
fonts, 237–238, 248
  attributes of, 211
  character sets for, 211
  for readability, 213–215
Food and Drug Administration (FDA), 314
footings in spreadsheets, 146
footnotes, 239, 260–262, 334
foreground, palette colors for, 272–273
foreign words, italics for, 220
formulas
  in tables and spreadsheets, 156–157
  for trends, 131–133
fractions, rules for, 222
functions for spreadsheets, 156, 162–163
fundamental analysis, 167, 178, 180, 187, 335, 349–350

G

gamut, color, 265–266
Gantt charts, 52–53, 73–75
Gantt, Henry, 74
geographic information, commas for, 216
geography
  cultural biases in, 54–57
  in sales charts, 58–59
grammar, 225, 246
grand totals in spreadsheets, 145

grayscale, 241, 255, 263, 269, 279–280, 341, 350

green, response to, 259

grids on radar charts, 84

growth (price appreciation), 178

H

hacking, 183

handouts, 280

HCV color system, 258

Head and Shoulders, 175–177, 180, 335, 350

headings
- aligning, 152
- in spreadsheets, 145

headlines, capitalization in, 221

healthcare professionals, color response by, 259–260

height of characters, 214

Helvetica font, 209

hidden font attribute, 210

high-low-close-open (HLCO) charts, 72–73, 172

highlighting data points, 48, 61, 64, 67, 130, 200

holiday colors, 269

horizontal bar charts, 76–77
- for Gantt charts, 52–53, 73–75
- for sales, 51
- for time, 51–52, 77

HSL color system, 276, 280–281

HSV color system, 258

hue, 257–258, 282

human faces, 60

hyphen (-), 198, 217, 235, 245, 251

I

illustrations, credit lines for, 207

importance, indicating
- in pie charts, 26–27, 29–30
- in *xy* charts, 81
- income (dividends), 178

increases, representation of, 44–45

independent variables in regression analysis, 140

Institute of Electrical and Electronics Engineers (IEEE), 233

intellectual property rights, 207

intensity in color, 257–258

international characters, 209

international signs, 45–46

interpolation, 16, 132–133, 169–170, 202, 320, 332, 337, 350

intersection of sets, 126

INTRATE function, 159, 161

investment decisions, 167, 177

investments
- functions for, 159–161
- trends in, 130–131

italics
- attribute for, 210–211
- rules for, 220

iterations for circular references, 161

J

Japan, radar charts in. *See* radar charts

JPEG (.jpg) file type, 192

justification in tables, 152

K

keys for series, 205, 262

keystroke error, 182, 198

keywords in table functions, 162

L

labels
- for axes, 65, 94–109
- for categories, 4–5
- for data points, 197–205
- for 100 percent bar charts, 40, 123–124
- for pie charts, 21–24, 22–23, 30–32
- for scales, 100–101, 204, 337
- in tables and spreadsheets, 156

legends
- color in, 205, 238, 337
- composing, 205
- for series, 205
- for tables, 150

legibility, 151, 212, 334

letterforms, 212, 351, 355

lighting with background color, 267–268

lightness in color, 257–258

line charts, 66
- baselines in, 68, 70
- stacked, 68
- for trends, 112–114, 129

linear regression, 140–141

linear scales, 103, 104

lines, slanting, 53–54

linked bar charts, 118–119

lists, 151
- in presentations, 228–229

Lo, Andrew W., 178

log-log scales, 105–106

logarithmic regression, 140–141

logarithmic scales, 103–106

logical factors, 180

logical table values, 163

lowercase letters, 209, 212–213, 246, 251, 338

M

MacKinlay, A. Craig, 178

magenta, response to, 259

major divisions on axes, 94

Malkiel, Burton G., 178

management with Gantt charts, 52–53, 73–75

manuals of style, 207, 223, 225, 232

maps, cultural biases in, 54–57

mathematicians' positioning rule, 24–25

maximum axis values, 95

McArthur's Universal Corrective Map of the World, 54–55

Mean Time Between Failures (MTBF), 300, 351

means, 11
- deviation from, 137–139

measurement units
- on axis titles, 202–203
- in labels, 4

measurements, 1–2

MEDIAN function, 162

medians, 13, 134–136

Mercator maps, 54–56

metadata analysis, 301–309, 342
metrics, 180, 318
middleware, 312
midpoints, 13
  in trends, 134–135
minimum axis values, 95
minor axis divisions, 94
mixed charts, 128
mixing data types in tables, 147–149
monochrome output, designing for, 279–280
monospace fonts, 211–213
moods, color for, 258–260, 268–269
movement for increases and decreases, 44–45
movie audiences, color response by, 258–259
moving averages, 14, 136–137
moving table information, 163–164
multivariate analysis, radar charts for, 84

N

names
  for database fields, 148, 282
  for table ranges, 155, 163–164
  Nanex, 171, 182, 184
NASDAQ, 178–179
natural laws, trend lines for, 133
natural logarithms, 140
neckline, 175–177, 301, 315
negative numbers, 158, 241
negatives, direction of, 44
nested parentheses, 163
Net Promoter Score (NPS), 297–300, 342, 352
New York Stock Exchange (NYSE), 172, 178–182 , 309
Non–Random Walk Down Wall Street, 178
numbered lists, 151
numbers
  aligning, 152
  averages of, 11–15
  categories for, 3
  labels for, 4–5
  precision of, 157–160
  spelling out, 222
  summaries of, 8–10
  trends in, 16

O

omissions
  ellipsis for, 271
  in pie charts, 35–36, 111
ordering pie chart slices, 24–26, 28–29
orders of magnitude, scales for, 103–106
ordinal values, 109
orientation
  of increases and decreases, 44–45
  in photographs, 44–45
  of pie chart slices, 24–26
  of scale labels, 204
  of time, 43–44
origins in Cartesian coordinate system, 62
outlier, 295, 308
outlines, 280
output media, converting, 269–270
overhead transparencies, 268
overlays in Cartesian coordinate system, 62–63
oversimplification in trends, 133–137

P

page margins, 249

page numbering, 250

paired–bar charts, 124–125

palettes, color, 265–266, 272–274, 280

Pantone Matching System (PMS), 280

parallel construction, 242, 352

paraphrasing, 207, 230

parentheses (), 163, 217–218, 241, 352
  in table arguments, 156

passive voice, 224–225

pastel colors, 269

patterns, 67, 133, 139, 167–187
  with color, 261–262
  for legends, 259
  in tables, 144

percentages
  100 percent charts for, 39–40, 123–124
  pie charts for. *See* pie charts

periods (.)
  in dates, 199
  with parentheses, 217–218
  with quotation marks, 219
  rules for, 218

permissions, 207

perspective in *xy* charts, 78–80

photographs
  color in, 262–263
  credit lines for, 207
  geographical biases in, 57
  orientation in, 44–45

picas, 209

pie charts, 19–21
  All Others category in, 35–36
  alternatives to, 39–41
  donut, 38
  importance of slices in, 26–27, 29–30
  labels for, 21–24, 30–32
  omissions in, 35–36, 111
  ordering slices in, 24–26, 28–29
  orientation of slices in, 24–26
  pie–column, 37
  for sales, 30–32
  two–pies, 32–34
  using, 109–111

pie-column charts, 37

pieces, pie chart. *See* slices, pie charts

plagiarism, 261

planar charts, 62

plotting styles, 108

PMS colors, 280–281

point charts, 65, 67

POM dependency, 114

population growth, 303–309

population standard deviation, 138

positioning
  pie chart slices, 24–26
  in *xy* charts, 81

possessive forms, 215

power regression, 140

PowerPoint, xv–xx, 51, 93, 190, 272, 317

precision
  in axis scales, 101
  in investment functions, 159–161
  in radar charts, 89–90

in tables, 157–160

prejudice. *See* cultural bias

prepositions, 221

printing tables, 151

priorities of Cartesian coordinate system objects, 62–63

programmed trading, 178–179, 182–185, 353

progress, direction of, 43–44

project management, Gantt charts for, 52–53, 73–75

projections, 32, 173–174, 180, 352, 353

proofreader's marks, 251

proofreading, 231–254

proportional representation in charts, 77, 190–191, 195, 202

proportionally spaced fonts, 211–213

proportions
- 100 percent charts for, 39–40, 123–124
- pie charts for. *See* pie charts

psychological factors, 180

psychology of color, 258–260

public domain material, 208, 353

pullback, 175–177, 184, 335, 353

punctuation, 198, 209, 214–221, 239–244

pure colors, 277

Q

quantities, 4
- in pie charts, 21–23, 30–32
- in *xy* charts, 47–49, 61, 64–65

question marks (?), 218–219

quotation marks (")
- rules for, 219
- in table functions, 162–163

quote stuffing, 182, 353

quoted material, capitalization in, 275

R

radar charts, 83
- characteristics of, 84
- example of, 85–87
- scales in, 89–92
- symmetry in, 87–88

Random Walk Down Wall Street, 178

range of color, 267

ranges
- of axes, 94
- bounds for, 95–98
- names for, 155, 163–164
- in tables and spreadsheets, 154–155

rate of change, emphasizing, 95–99

rate of return functions, accuracy of, 159–161

ratios
- 100 percent charts for, 39–40, 123–124
- pie charts for. *See* pie charts

raw data
- processing, 2
- scatter charts for, 67

read-ahead problem, 228

readability
- contrast for, 267

readability index, 225, 354

reading level, 225

reasonableness of table results, 164

records in database tables, 148

red, response to, 259

reduction, data. *See* data reduction

regression analysis, 139–142

relative amounts
- 100 percent charts for, 39–40, 123–124
- pie charts for. *See* pie charts

rendering tool, color as, 262–263

reporting periods, 237

reports, 231–254

reports, tables in, 153

resistance line, 173–174, 176, 186, 354

RGB color method, 258, 275–276, 280–281

rolling release, 294, 354

root cause, 303

rotating
- axes, 75–77
- pie chart slices, 29

ROUND function, 158, 161

rounding
- in radar charts, 90
- in tables, 158, 161

rows in tables, 143, 145, 148, 154

S

sales
- geographical representations of, 58–59
- horizontal bar charts for, 51
- pie charts for, 30–32
- two-pie comparisons for, 32–34

sample standard deviation, 138

sans-serif fonts, 266

saturation, color, 257–258

scalable fonts, 264

scale break, 194, 354

scale factors for axes, 100–101

scales, 94
- bounds for, 94–98
- for dual y–axis charts, 101–103
- labels for, 100–101, 258
- logarithmic, 103–106
- in radar charts, 89–92

scatter charts, 65, 67, 256

scatter–line charts, 128

schedule management, Gantt charts for, 52–53, 73–75

schemes, color, 265–266, 272

screen
- colors on, 277–278
- readability on, 266–268

secondary hues, 282

sectors, pie chart. *See* slices, pie charts

Securities and Exchange Commission (SEC), 181

segmented bar graphs, 9–10

segments, pie chart. *See* slices, pie charts

selecting
- charts, 107–109
- colors, 266–272
- fonts, 265–266

semi–logarithmic scales, 103–104

semicolon (;), 219–220, 244

sense, reviewing for, 226

serial comma rule, 216, 233, 355
serial date values, 162
series, 1, 65, 67
  in dual y–axis charts, 102
  legends for, 102, 203, 205
  for tables, 150
serif fonts, 212, 338
sets, Venn diagrams for, 125–128
shade, 257–258
shapes in radar charts, 84, 87–88
short-sale order, 177–178, 355
signposts, 151, 227, 230, 338–339, 355
signs, international, 45–46
single quotation marks (’), 219
size
  in pie chart slice ordering, 28
  in two–pie comparisons, 32–34
skewed averages, 12
skip factors for scale labels, 204, 355
slanting lines, 53–54
slanting scale labels, 204
slices, pie charts
  exploding, 27, 29, 37, 108
  importance of, 26–27, 29–30
  labels for, 21–24
  ordering, 24–26, 28–29
  orienting, 24–26
slip cards, 268
slippage on Gantt charts, 73
slope of lines, 129–130, 173–174, 190–191, 348, 355
small caps font attribute, 210–211
smart quotes, 246, 355
software applications for business charts, 51
special words and phrases, italics for, 221
spelling, 223, 234–235
spoofing, 183, 355
spreadsheets. *See* tables and spreadsheets
stacked charts, 68–71, 122
  area charts, 40–41, 115–117
  bar charts, 9–10, 39, 120–121
  vs. cumulative, 71, 122
staggered scale labels, 204
Standard & Poor’s 500 (S&P 500), 172
standard deviation, 138
starting points for *y*-axes, 98–99
stocks, charts for, 72–73, 167–188
Stoenner, Hans-Martin, 303–309
stop-loss order, 174, 177–178, 181, 355
strikethrough font attribute, 210–211, 338
Structured Query Language (SQL), 312
style, 231–254
style sheets, 270, 276, 355
subscripted variables, 95
subsets, 126
substantial similarity standard, 207, 353, 355
subtitles, color for, 266–267
subtotals in spreadsheets, 145–146
subtractive color process, 280
subtypes, 108
SUM function, 156
summaries, 8–10, 150–151

support lines, 173, 176–178, 181, 186, 353, 356

symbols, 45–46, 67, 75, 77, 209, 241, 246–248, 251

symbols in *xy* charts, 77–78

symmetry in radar charts, 87–88

T

tables and spreadsheets, 143–144
- alignment in, 152
- cells in, 154
- charts from, 165
- circular references in, 160–162
- columns and rows in, 143, 145, 148, 151, 154
- fonts for, 265, 267
- formulas in, 156–157
- functions for, 162–163
- mixing data types in, 147–149
- patterns in, 144
- precision in, 157–160
- printing, 151
- ranges in, 154–155
- reasonableness of results in, 164
- in reports, 153
- rules for, 150–152, 161–165
- styles for, 145–149

task management, Gantt charts for, 52–53, 73–75

TBILLYIELD function, 159, 161

technical analysis, 72, 167–187, 335, 356

templates, 248, 270, 272, 276, 278, 317, 355, 356

testing tables, 164

text
- aligning, 152
- color for, 266–268
- palette for, 272–273

thematic color coding, 264

thousands separator, 240

ticks
- on axes, 94
- on HLCO charts, 72

time
- direction of, 43–44
- in horizontal bar charts, 51–52, 77
- in *xy* charts, 47–49, 61, 64

time analysis, Gantt charts for, 52–53, 73–75

time series, 304, 343, 356

Times New Roman font, 211

tint, 257–258

title case, 243

title screens, 338

titles (articles and headings)
- axis, 202–203
- color for, 266–268
- for pie charts, 22

titles (persons), capitalizing, 222

tone in color, 257–258

totals, 8–10
- cumulative bar charts for, 122–123
- in spreadsheets, 145–147

translations, copyright protection for, 208

trends, 16, 129–131, 174, 184–185, 303–309
- area charts for, 115
- formulas for, 131–133
- line charts for, 112–114
- mean deviations in, 137–139
- metadata, 301–309

- oversimplification in, 133–137
- regression analysis for, 139–142
- troughs, 177, 352, 356
- two-pie comparisons for, 32–34
- volume trends, 176

TRUE table values, 163

TRUNC function, 158–159, 161–162

truncation, 240, 320
- in radar charts, 90
- in tables, 158–159, 161–162

two-pie comparisons, 32–34

typefaces, 208–211, 256, 337, 349, 356

types of charts, 107–128

U

uncertainty, error bars for, 138–139

underlining attribute, 211

unintended consequences, 314–317

union of sets, 126–127

units
- abbreviations for, 277
- on axis titles, 100–102
- in labels, 4, 222–223

University of Chicago, 225, 232

unprocessed data, 2

unseen hand, 186

V

values, 4
- in pie charts, 21–23, 30–32
- in *xy* charts, 47–49, 61, 64–65

vanishing points in perspective, 78–80

variables in regression analysis, 140

vector graphics, 192, 357

Venn diagrams, 125–128

vertical bar charts
- clustered, 120–121
- cumulative, 122–123
- 100 percent, 118–120, 123–124
- stacked, 120–121

vertical scale labels, 258

video, color for, 268–269

viewing distance, 266–267

viewport, 189–195, 336

visual cueing, 230, 357

volatility, 129, 193, 304–306, 309, 342, 351, 357

volume bars, 176, 357

volume (measurement), area charts for, 114–118

volume (sales) on HLCO charts, 73, 176, 357

W

Wall Street Journal, 183

“we,” use of, 224

wedges, pie chart. *See* slices, pie charts

weighting
- in radar charts, 91
- in regression analysis, 140

Windows Metafile (.wmf) file type, 192

work projects, Gantt charts for, 52–53, 73–75

workpapers, 145

X

x-axes, 48–50, 61, 64–65

xy charts, 47–48, 61–82, 88, 93–106, 112–125, 130, 142, 192, 202, 324–328
- cultural biases in, 47–53
- cumulative, 71–72
- Gantt charts, 73–75
- high–low–close–open, 72–73
- perspective in, 78–80
- position in, 81
- rotating axes on, 75–77
- scaling, 100, 189–194, 203
- stacked, 68–70
- symbols on, 77–78
- trends in. *See* trends

Y

y–axes, 48–50, 61, 64–65
- bounds for, 95–98
- dual, 101–103, 113–114
- starting points for, 98–99

year-end results, 237

years, types of, 237, 339

yellow, response to, 259

YIELD function, 159, 161

Z

z–axis, 62–63

zero starting points for y–axes, 98–99

zeros
- in averages, 15
- division by, 162

About the Author



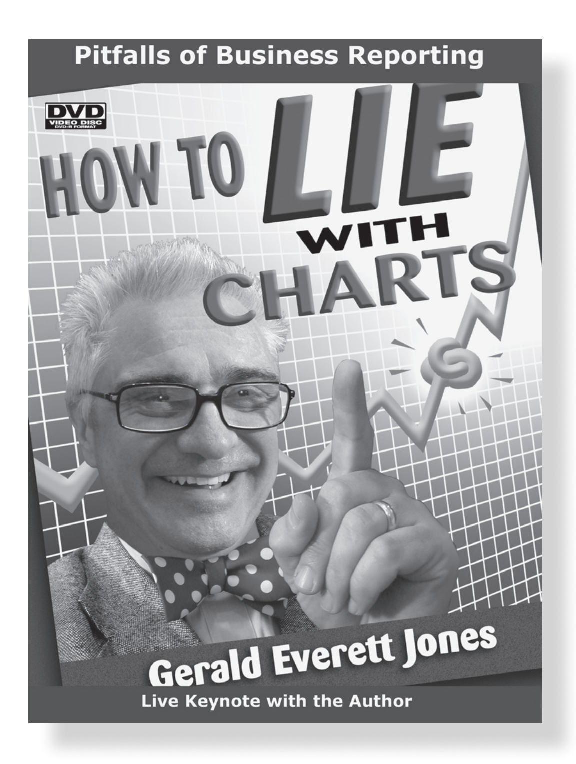
Gabriella Muttone Photography

Gerald Everett Jones has managed major proposal and business analysis projects in government, healthcare, and commercial sectors. He is the author of more than 25 business and technical books on digital media production, including numerous how-to books on legacy charting applications such as Harvard Graphics and Freelance Graphics. He has also written extensively on PowerPoint and Excel. He has hands-on background in IT systems development and was project leader on the ARTIS computer graphics system, a precursor of PowerPoint. He often leads seminars on the topic of achieving honesty and clarity in business presentations.

INSTRUCTIONAL VIDEOS

Pitfalls of Business Reporting Series

The material in *How to Lie with Charts: Investor's Edition* is available in three separate videos from LaPuerta Books and Media. Visit **www.lapuerta.tv/book-catalog** to purchase them as DVDs or streaming Amazon Video (buy or rent).



How to Lie with Charts (Second Edition)

Advice for Analysts and Investors

48 minutes, NTSC 4:3 UPC: 885444435683

This live keynote by the author covers most of the content of the Second Edition. Learn how to spot misinformation in visual presentations—especially your own. If you're using a computer to generate charts for meetings and reports, you don't have to be taught how to lie—you're already doing it. In this interactive seminar presented at an Operational Excellence conference of a major corporation, the author uses wit and lucid illustrations to help business professionals and investors avoid getting tricked by what they see.

- More than 100 examples
- Nontechnical and entertaining
- Also previews this Investor's Edition of the text
- Bonus how-to slideshow on the DVD



LaPuerta

www.lapuerta.tv

Santa Monica, California

The LaPuerta imprint and its logo—an open door—symbolize unlimited access to knowledge, opportunity, innovation, fascination, and delight.



Technical Analysis Basics

54 minutes, NTSC 16:9 UPC: 887936546778

This live seminar applies the principles in *How to Lie with Charts* to the charting of financial results. Specifically, it focuses on the new material presented in Chapters 10 (Technical Analysis Basics) and 11 (Viewports and Dashboards) of the Investor's Edition. The author's focus is on market technical analysis, but the wider implications of visual perception and interpretation bear on all aspects of creating clear communications, including proposal and presentation graphics.

To apply one of Murphy's Laws, "Technical analysis is a systematic method of coming to the wrong conclusion with high confidence." You'll see how seemingly arbitrary choices about parameters and presentation modes can make all types of charts difficult to understand or even misleading. Relevant topics include separating signals from noise, axis manipulation, scaling, 3D effects, statistical analysis, and the human tendency to see meaning in patterns.

Jones not only simplifies technical analysis but also places it in perspective as a tool among various approaches to making decisions about what's happening on a daily basis in stock and commodities markets.

Financial Proofreading

29 minutes, NTSC 16:9 UPC: 887936547782

Learn how to make your financial reports both professional and classy. Here's the new material from Chapter 13 (Financial Proofreading) of the Investor's Edition.

Diligent proofing in itself won't make a killer impression. But failing to proof for both accuracy and consistency could have woeful results.

In this narrated slideshow seminar, the author brings the point home with a simple formula:

Inconsistency
= Perceived Sloppiness
= Lack of Thoroughness
= Mistrust of Your Results!

In a brief, half-hour session, you'll get a comprehensive survey of: variations in professional style rules and how to enforce consistency, along with guidelines for footnotes, punctuation, spelling, appearance, proofing methods, and even tips for composing email.

Perhaps numbers don't lie, but you can't afford a comma or a decimal point in the wrong place!

SUITABLE FOR YOUNG ADULTS

Also from ***LaPuerta Books and Media*** *www.lapuerta.tv/book-catalog*

by ***Gerald Everett Jones***

Mr. Ballpoint

ISBN: 978-0-9856227-4-9 hardcover;
978-0-9856227-5-6 Kindle

Humorous Fiction (ages 14 and up)

In 1945, Milton Reynolds introduced the ballpoint pen to the United States and triggered the biggest single-day shopping riot in history. Reynolds, an exuberant huckster who had already made and lost several fortunes, again became an overnight millionaire and then bragged that he "stole it fair and square."

Milton was a man ideally suited to his time—the post-war boom when the salesman was king and all of the rules had yet to be written. He was an old-fashioned, silver-tongued American peddler who would do almost anything—ethical or otherwise—to close a deal. His son Jim was a quiet Boy Scout who couldn't tell a lie—even when he needed to.

Ballpoint is a father-son relationship story, told from Jim's point of view, about coping with Milton's outrageous schemes, then their sudden success. The conflicts between these two fundamentally different characters drive the comedy of the story.

"Forgotten history brought to life. If you ever wanted to know how to play the game of life and have a blast doing it, read *Mr. Ballpoint*. Perfect for our library and book clubs."

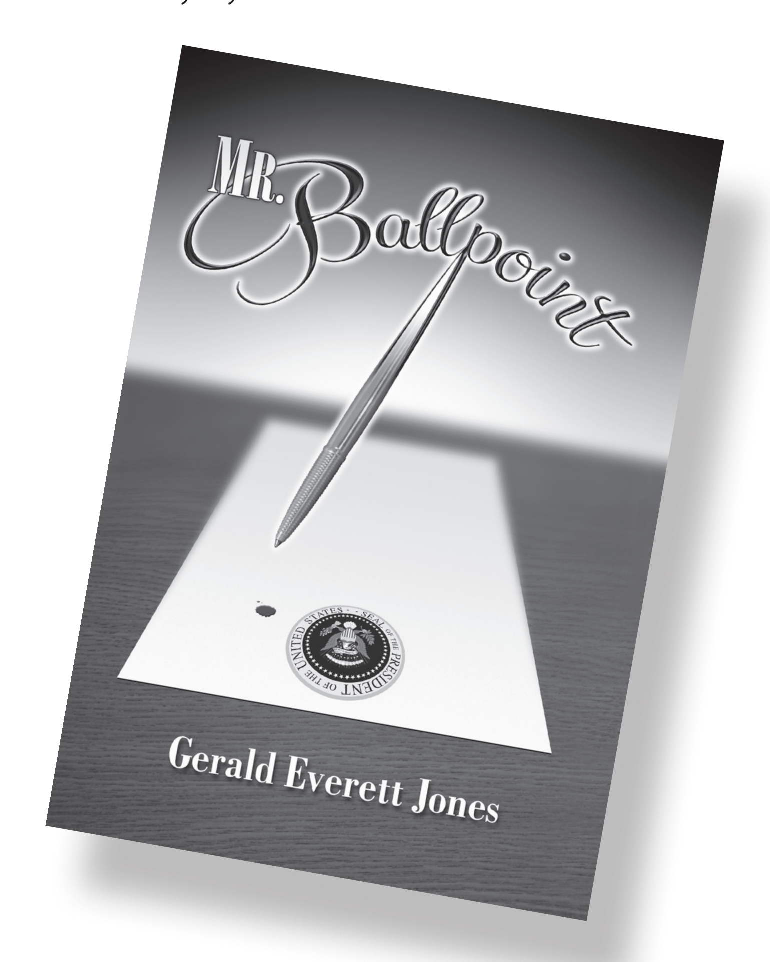
—**Deborah Vaden**, Manager of Libraries, City of Irving, Texas

"Gerald Everett Jones certainly knows how to tell a story. Engaging, funny, and full of the cross-genre capability that marks all of Jones's novels, *Mr. Ballpoint* will surprise and delight readers of all ages."

—**Magdalena Ball**, host of The Compulsive Reader and author of *Black Cow* and *Sleep Before Evening*

"If only I had known about this marvelous invention before I started my writing career! No typewriter! No computer! The story calls to mind the old adage about pioneers and arrows. I salute Mr. Jones for his delightful and insightful tale."

—**Marvin J. Wolf**, author of *Rotten Apples: Tales of New York Crime and Mystery* and *For Whom The Shofar Blows*



"In the ever-expanding universe of Gerald Everett Jones, this is a funny, typically quirky and engrossing read. Milton Reynolds, a rollicking hustler who harvested gravity, is a bigger-than-life character with Best Actor written all over it. Should be a movie by Monday."

—**Morrie Ruvinsky**, author of *Meeting God or Something Like It* and *Misfits of Science*